Load Balancing with Almost Blank Subframe Control in Heterogeneous Cellular Networks

Doo-hyun Sung and John S. Baras
Institute for Systems Research and Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
Email: dhsung@umd.edu, baras@isr.umd.edu

***Abstract*—In heterogeneous cellular networks, the network capacity can be significantly enhanced via user offloading between macro- and pico-tier. Due to the strong cross-tier interference from macrocells to picocells in co-channel deployment, the user offloading needs to be jointly optimized with macrocells' transmit power nulling, i.e., almost blank subframes (ABSs) in 3GPP LTE-A systems. In this paper, we discuss the network-wide utility maximization problem where the cell association and the number of ABSs are jointly optimized. Due to the NP-hardness of the formulated problem, an online heuristic algorithm is proposed where (i) user load is re-distributed based on users' expected data rate by handover for the current number of ABSs, and (ii) the current number of ABSs is changed (increased or decreased) by estimating the possible users to be offloaded based on users' expected data rate by handover and ABS change in a gradient descent manner. Through simulations, we demonstrate that the proposed algorithm not only improves the average data rate of users, but also achieves better fairness among users.**

I. Introduction

Heterogeneous cellular networks (HetNets) have been considered as one of the key features to cope with the increasing mobile data traffic by deploying low-power small cells such as picos, femtos, and relays inside traditional macrocells' coverage to achieve spatial cell-splitting gain [1], [2].

In the 3rd Generation Partnership Project Long-Term Evolution Advanced (3GPP LTE-A) systems, mobile stations (MSs) are associated with a BS with the strongest received pilot signal after applying a positive offset to the received power from small cells, which is known as cell range expansion (CRE). To mitigate the cross-tier interference from macro BSs (MBSs) to pico MSs (PMSs) in co-channel HetNets, the time-domain transmit power nulling at MBSs, known as almost blank subframes (ABSs), has been proposed. During configured ABSs, MBSs are not transmitting any signal except essential ones for system maintenance and backward compatibility such as cell-specific reference signals, synchronization signals, and broadcast system information signals.

Even though macro MS (MMS) offloading toward pico BSs (PBSs) can be achieved by the CRE operation, MS load balancing is still a challenge in HetNets as the received signal strength-based cell association is hard to cope with load imbalance. Moreover, the number of ABSs needs to be jointly optimized to compensate strong cross-tier interference to PMSs located in the expanded range.

There have been several researches on load balancing for general multi-cell wireless networks [3], [4] and for heterogeneous cellular networks [5], [6]. However, there have been a few studies on joint optimization of the cell association and the number of ABSs. In [7], authors discuss this joint optimization problem and transform the combinatorial problem into a convex form by relaxing a binary cell association. Although the solution provides an upper-bound of performance, this offline approach and MSs' multiple associations with BSs are less viable in practice.

In this paper, we propose a practical online algorithm to solve the joint optimization problem of the cell association and the number of ABSs. The proposed algorithm consists of two stages - load balancing and ABS control. For a given number of ABSs t, the load balancing stage is first performed where MS handovers are triggered such that the network-wide utility is improved by handovers. When there is no MS available for handover, i.e., MS load is balanced under the current ABS duration t, the ABS control stage is then performed where the possible MSs for further offloading (macro $\leftrightarrow$ pico) are estimated by changing the ABS duration ($t \rightarrow t \pm 1$).

II. System Model

The network model considered in this paper is a heterogeneous downlink cellular network consisting of two tiers - MBSs and PBSs overlaid within the MBSs' coverage. Based on Table I, we derive two expressions of average received signal-to-interference plus noise ratio (SINR) in non-ABSs and

TABLE I
List of parameters and variables

Notation	*Description*
$\mathcal{U}$	Set of MSs
$\mathcal{U}_b$	Set of MSs associated with BS b
$\mathcal{B}$	Set of BSs ($= \mathcal{B}_m \cup \mathcal{B}_p$)
$\mathcal{B}_m$	Set of MBSs
$\mathcal{B}_p$	Set of PBSs
x_{ub}	Association indicator of MS u with BS b
r_u	Expected average throughput of MS u
$r_{ub}(t)$	Long-term average rate of MS u from BS b with t ABSs
c_{ub}	Achievable link rate of MS u from BS b in non-ABSs
$\bar{c}_{ub}$	Achievable link rate of MS u from BS b in ABSs
W	System bandwidth
T	ABS periodicity in subframes

978-1-4799-3512-3/14/$31.00

ABSs as follows. The average SINR in non-ABSs at MS u from BS b, denoted by Γ_{ub}, is expressed as

$$\Gamma_{ub} = \frac{P_b h_{ub}}{\sum_{b' \in \mathcal{B}, b' \neq b} P_{b'} h_{ub'} + \sigma^2}, \quad (1)$$

where P_b, h_{ub}, and σ^2 denote the transmit power of BS b, the average channel gain of a link between MS u and BS b including path loss, shadowing, and fast fading, and the power of additive white Gaussian noise, respectively. In ABSs, all MBSs' transmit power is set to be zero. Hence, the average SINR in ABSs at MS u from BS b, denoted by $\bar{\Gamma}_{ub}$, is expressed as

$$\bar{\Gamma}_{ub} = \begin{cases} \dfrac{P_b h_{ub}}{\sum_{b' \in \mathcal{B}_p, b' \neq b} P_{b'} h_{ub'} + \sigma^2} & \text{if } b \in \mathcal{B}_p \\ 0 & \text{if } b \in \mathcal{B}_m. \end{cases} \quad (2)$$

Given Γ_{ub} and $\bar{\Gamma}_{ub}$, two types of achievable link rates, c_{ub} in non-ABSs and $\bar{c}_{ub}$ in ABSs, are derived using Shannon's formula as

$$\begin{aligned} c_{ub} &= W \log_2 (1 + \Gamma_{ub}) \\ \bar{c}_{ub} &= W \log_2 \left(1 + \bar{\Gamma}_{ub}\right) \end{aligned} \quad \forall u \in \mathcal{U} \quad \forall b \in \mathcal{B}, \quad (3)$$

respectively, where W is the system bandwidth. For any MBS $b \in \mathcal{B}_m$, the average link rate $\bar{c}_{ub}$ in ABSs becomes zero as $\bar{\Gamma}_{ub}$ becomes zero.

$$P_b = 0 \ \rightarrow \ \bar{\Gamma}_{ub} = 0 \ \& \ \bar{c}_{ub} = 0 \quad \forall b \in \mathcal{B}_m \quad \forall u \in \mathcal{U}_b. \quad (4)$$

III. Problem Formulation

In this section, we formulate an optimization problem in which the objective is to maximize the network-wide utility by configuring both the user association with cells in a load-distributed manner and the number of ABSs to improve this load distribution between MBSs and PBSs as follows:

$$\begin{aligned} \max_{\underline{x}, t} \sum_{u \in \mathcal{U}} U(r_u) &= \max_{\underline{x}, t} \sum_{u \in \mathcal{U}} U \left(\sum_{b \in \mathcal{B}} x_{ub} r_{ub}(t) \right) \\ &= \max_{\underline{x}, t} \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} x_{ub} \, U\left(r_{ub}(t)\right) \end{aligned} \quad (5)$$

where $U(\cdot)$ is an increasing, strictly concave, and continuously differentiable utility function, $\underline{x}$ is an association indicator vector $\{x_{ub} : u \in \mathcal{U}, b \in \mathcal{B}\}$ representing whether MS u is associated with BS b $(= 1)$ or not $(= 0)$ and satisfying $\sum_{b \in \mathcal{B}} x_{ub} = 1$ for all u, and t is the number of ABSs configured at MBSs. We assume synchronous ABS operation where all MBSs follow the same ABS configuration such as the periodicity, the start offset, and the duration.

For the utility function $U(\cdot)$, we utilize the log utility function $U(r) = \log(r)$ as the previous work in [8] has shown that proportional fairness among users could be achieved when the sum of logarithmic utilities is maximized.

To derive the long-term average rate $r_{ub}(t)$, we assume that the proportional fairness scheduler is used. Following the long-term behavior of the proportional fairness scheduler in [9], $r_{ub}(t)$ can be expressed as

$$r_{ub}(t) = \frac{G(K_b) \left((1 - \frac{t}{T}) c_{ub} + \frac{t}{T} \bar{c}_{ub} \right)}{K_b}, \quad (6)$$

where K_b is the number of MSs associated with BS b which is derived as $K_b = \sum_{u \in \mathcal{U}} x_{ub}$ and $G(\cdot)$ denotes a multi-user diversity gain which can be calculated as $G(k) = \sum_{i=1}^{k} \frac{1}{i}$. Following the configured ABS duration t, $(T - t)$ non-ABSs and t ABSs are basically available for each MS.

By applying the logarithmic utility function and plugging (6) to (5), we have the formulated optimization problem as follows:

$$\max_{\underline{x}, t} \sum_{u \in \mathcal{U}} \sum_{b \in \mathcal{B}} x_{ub} \, \log \left(\frac{G(K_b) \left((1 - \frac{t}{T}) c_{ub} + \frac{t}{T} \bar{c}_{ub} \right)}{K_b} \right) \quad (7a)$$

$$\text{s.t.} \quad x_{ub} = \{0, 1\} \quad \forall u \in \mathcal{U} \quad \forall b \in \mathcal{B} \quad (7b)$$

$$\sum_{b \in \mathcal{B}} x_{ub} = 1 \quad \forall u \in \mathcal{U}, \quad (7c)$$

$$t = \{0, 1, \cdots, T - 1\}. \quad (7d)$$

As discussed in previous work [3], [4], the cell association problem, i.e., the optimization problem (7) with a fixed ABS duration t, is a 0-1 knapsack problem, therefore it is NP-hard. The authors in [3] present an *offline* algorithm which can obtain the optimal cell association in a polynomial time by fixing the number of associated MSs with each BS. For every K_b configuration, the cell association problem is equivalent to the maximum weighted matching problem. This offline algorithm, however, has computational complexity of $O\left(|\mathcal{U}|^{|\mathcal{B}|+3/2}\right)$ which could be too complex in heterogeneous cellular networks as the number of BSs in heterogeneous cellular networks is much larger than that in traditional cellular networks. Moreover, in our problem formulation, the ABS duration t also needs to be jointly optimized along with the cell association. Therefore, we develop an *online* heuristic algorithm inspired by [4] in the next section.

IV. Proposed Algorithm

The main motivation of load-aware cell association in general is that selecting the serving BS with the strongest received signal strength from it doesn't necessarily mean that MSs can achieve the highest average rate because the average rate depends on both the received signal strength and the user load shown in (6). In the CRE-enabled scenario, the interference would be larger than the desired signal for PMSs which severely degrades their average rate. Thus, the use of ABSs plays an important role as it can change MSs' average rate changes depending on their associated tiers. This could be used to trigger MSs' tier selection between macros and picos. Based on these observations, we develop the following properties which are crucial for our algorithm design.

Proposition 1. (Condition for MS handover under ABS *t*) Assume MS u is associated with BS b with ABS duration t and the number of users in BS b and BS $b^{'}$ are large. Then transferring the MS u from BS b to BS $b^{'}$ improves the network-wide utility if

$$\log \frac{G(K_{b'} + 1) e_{ub'}(t)}{K_{b'} + 1} - \log \frac{G(K_b) e_{ub}(t)}{K_b} > \delta_{bb'}^{HO} \quad (8)$$

$$\sum_{i\in\mathcal{B}_m}\sum_{u\in\mathcal{U}_i^{OL}}\left[\log\frac{G(K_{b_u^p}+m_{b_u^p})e_{ub_u^p}(t+1)}{K_{b_u^p}+m_{b_u^p}}-\log\frac{G(K_i)e_{ui}(t)}{K_i}\right]-\sum_{i\in\mathcal{B}_m}\left[n_i\left(\frac{1-\frac{n_i}{K_i}}{\gamma+\log K_i}-\frac{1}{T-t}\right)+\frac{K_i}{T-t}\right]$$
$$+\sum_{j\in\mathcal{B}_p}\left[\frac{m_j}{\gamma+\log K_j}+\sum_{v\in\mathcal{U}_j}\frac{a_{vj}-1}{T+t(a_{vj}-1)}\right]\ >\ 0 \tag{10}$$

$$\sum_{j\in\mathcal{B}_p}\sum_{u\in\mathcal{U}_j^{OL}}\left[\log\frac{G(K_{b_u^m}+n_{b_u^m})e_{ub_u^m}(t-1)}{K_{b_u^m}+n_{b_u^m}}-\log\frac{G(K_j)e_{uj}(t)}{K_j}\right]-\sum_{j\in\mathcal{B}_p}\left[\frac{m_j\left(1-\frac{m_j}{K_j}\right)}{\gamma+\log K_j}+\sum_{v\in\mathcal{U}_j\setminus\mathcal{U}_j^{OL}}\frac{a_{vj}-1}{T+t(a_{vj}-1)}\right]$$
$$+\sum_{i\in\mathcal{B}_m}\left[\frac{n_i}{\gamma+\log K_i}+\frac{K_i}{T-t}\right]\ >\ 0 \tag{12}$$

where $e_{ub}(t)=(1-\frac{t}{T})c_{ub}+\frac{t}{T}\bar{c}_{ub}$. $\delta_{bb'}^{HO}$ represents the net utility change between BS b and b' which is expressed as

$$\delta_{bb'}^{HO}=\frac{1-\frac{1}{K_b}}{\gamma+\log K_b}-\frac{1}{\gamma+\log K_{b'}} \tag{9}$$

where γ is the Euler-Mascheroni constant ($=0.5772\cdots$).

Proof: Refer to Appendix A.

Proposition 2. (Condition for ABS increment from *t* to *t*+1) Assume MBS $i\in\mathcal{B}_m$ selects a subset of its associated MMSs to be offloaded to PBSs, denoted by $\mathcal{U}_i^{OL}$ $\left(|\mathcal{U}_i^{OL}|=n_i,\ n_i\ll K_i\right)$ and PBS $j\in\mathcal{B}_p$ accommodates m_j MMSs out of overall MMSs to be offloaded $(m_j\ll K_j)$, i.e., $\sum_{i\in\mathcal{B}_m}n_i=\sum_{j\in\mathcal{B}_p}m_j$. Then ABS increment from t to $t+1$ with offloading MMSs improves the network-wide utility (i.e., the net utility change $\Delta U_+^{ABS}(t+1)>0$) if the condition in (10) is met, where b_u^p is the target PBS to which the MMS u is handed over, $m_{b_u^p}$ is the total number of MMSs that the target PBS b_u^p would accommodate, and a_{vj} is the ratio of the achievable link rate at PMS v with PBS j in ABSs to that in non-ABSs $(=\bar{c}_{vj}/c_{vj})$.

Proof: Refer to Appendix B.

Remark: Three terms in (10) can be represented respectively as follows:

$$\Delta U_m^{OL}(t+1)+\Delta U_m(t+1)+\Delta U_p(t+1)>0, \tag{11}$$

where $\Delta U_m^{OL}(t+1)$, $\Delta U_m(t+1)$, and $\Delta U_p(t+1)$ denote the expected net utility change of MMSs to be offloaded to PBSs, MMSs remaining in MBSs, and PMSs by increasing the ABS duration from t to $t+1$, respectively.

Proposition 3. (Condition for ABS decrement from *t* to *t*-1) Assume PBS $j\in\mathcal{B}_p$ selects a subset of its associated PMSs to be offloaded to MBSs, denoted by $\mathcal{U}_j^{OL}$ $\left(|\mathcal{U}_j^{OL}|=m_j,\ m_j\ll K_j\right)$ and MBS $i\in\mathcal{B}_m$ accommodates n_i MMSs out of overall PMSs to be offloaded $(n_i\ll K_i)$, i.e., $\sum_{j\in\mathcal{B}_p}m_j=\sum_{i\in\mathcal{B}_m}n_i$. Then ABS decrement from t to $t-1$ with offloading PMSs improves the network-wide utility (i.e., the net utility change $\Delta U_-^{ABS}(t-1)>0$) if the condition in (12) is met, where b_u^m is the target MBS to which the PMS u is handed over, $n_{b_u^m}$ is the total number of PMSs that the target MBS b_u^m would accommodate.

Proof: Refer to Appendix B.

Remark: Three terms in (12) can be represented respectively as follows:

$$\Delta U_p^{OL}(t-1)+\Delta U_p(t-1)+\Delta U_m(t-1)>0, \tag{13}$$

where $\Delta U_p^{OL}(t-1)$, $\Delta U_p(t-1)$, and $\Delta U_m(t-1)$ denote the expected net utility change of PMSs to be offloaded to MBSs, PMSs remaining in PBSs, and MMSs by decreasing the ABS duration from t to $t-1$, respectively.

From Eq. (10), (12), we observe that the net utility change of MBSs and PBSs by ABS control, apart from that of MSs to be offloaded, can be represented in a simple form.

Based on the observations discussed above, we describe how the proposed algorithm optimizes the cell association and the ABS duration. Considering the required procedure and signaling for measurement and reporting, the joint optimization is divided into two stages - the MS load-balancing and the ABS control. In the MS load-balancing stage, MS handovers are performed among BSs in a way that the network-wide utility is increased in a gradient-descent manner under the current ABS duration t. In the ABS control stage, the ABS increment (+1) or decrement (-1) is examined by estimating the possible MSs that can be offloaded based on MSs' measurement reports and corresponding net utility changes $U_+^{ABS}(t+1)$ and $U_-^{ABS}(t-1)$.

A. Stage 1: MS Load-balancing under ABS Duration t

By neighbor cell measurement and user load information, every MS u calculates the expected data rates of its neighboring cells by handover, and reports to the serving BS b_u the best target BS b_u' from which it can achieve the largest logarithmic ratio $\phi_u(t)$ by handover as

$$\phi_u(t)=\log\frac{G(K_{b_u'}+1)e_{ub_u'}(t)}{K_{b_u'}+1}\frac{K_{b_u}}{G(K_{b_u})e_{ub_u}(t)}\quad\forall u\in\mathcal{U}. \tag{14}$$

Suppose each BS reports their best candidate MS to the central coordinating entity, the central entity chooses the best MS u^* that achieves the largest utility increment by this handover,

$$u^* = \arg\max_{u \in \mathcal{U}^{HO}} \left(\phi_u(t) - \delta^{HO}_{b_u b'_u} \right), \quad (15)$$

where $\mathcal{U}^{HO}$ is the set of candidate MSs selected by each BS and $\delta^{HO}_{b_u b'_u}$ can be calculated based on (9). As discussed in (8), the handover can be done only if the selected MS u^* satisfies the following condition with a hysteresis margin $\delta^{HO}_h > 0$ to prevent possible ping-pong effects:

$$\phi_{u^*}(t) - \delta^{HO}_{b_{u^*} b'_{u^*}} > \delta^{HO}_h. \quad (16)$$

For a given ABS duration t, the load-balancing operation is performed until there is no MS that satisfies the condition in (16) to improve the network-wide utility by handover.

B. Stage 2: ABS Control from t to $t \pm 1$

By neighbor cell measurement and user load information, every MS u calculates expected data rates with its neighboring BSs by handover along with ABS change (ABS increment by +1 and ABS decrement by -1). Then, the MS reports the best target BS b'_u and the best ABS value t_u (either +1 or -1) to the serving BS b_u for which it can achieve the highest ratio $\phi_u(t + t_u)$ by handover along with ABS change as

$$\begin{aligned} &\phi_u(t + t_u) \\ &= \log \frac{G(K_{b'_u} + 1) e_{ub'_u}(t + t_u)}{K_{b'_u} + 1} \frac{K_{b_u}}{G(K_{b_u}) e_{ub_u}(t)} \quad \forall u \in \mathcal{U}, \end{aligned} \quad (17)$$

where $\phi_u(t + t_u) > 0$ means that the MS can be considered as a candidate MS for offloading with ABS change t_u.

Unlike load-balancing in stage 1 where a single MS is handed over to a target BS regardless of the BS type, ABS control only considers multiple MSs offloading from MBSs to PBSs, or vice versa. For instance, when the ABS duration increases this leads to MMSs' data rate degradation as the number of non-ABSs decreases. Hence, a certain number of MMSs should be offloaded to PBSs to compensate the possible MMSs' throughput degradation. When the ABS duration decreases, on the other hand, this leads to PMSs' data rate degradation as the number of ABSs decreases. Therefore, a certain number of PMSs needs to be offloaded to MBSs. As a result, we focus on two cases - ABS increment by +1 with offloading MMSs to PBSs and ABS decrement by -1 with offloading PMSs to MBSs.

After receiving MSs' measurement report messages, every MBS reports to the central coordinating entity the number of currently associated MMSs and information of candidate MMSs for offloading to PBSs such as the target PBS id and $\phi_u(t + 1)$. In case of PBSs, each PBS reports the number of currently associated PMSs, the net utility change of them by ABS increment calculated as

$$\sum_{u \in \mathcal{U}_j} \frac{a_{uj} - 1}{T + t(a_{uj} - 1)} \quad \forall j \in \mathcal{B}_p, \quad (18)$$

and information of candidate PMSs for offloading to MBSs such as the target MBS id, $\phi_u(t - 1)$, and $\frac{a_{uj} - 1}{T + t(a_{uj} - 1)}$. As observed in (10), (12), the net utility change by ABS decrement can be obtained by changing the sign of that of ABS increment and subtracting that of PMSs to be offloaded.

The central coordinating entity examines if the current ABS duration t needs to be changed by $+1$ or -1 via the backhaul messages from BSs. To check if the condition in (10), (12) is satisfied, the central coordinating entity needs to determine the number of offloading MMSs for ABS increment and offloading PMSs for ABS decrement.

Since MSs' reported ratio $\phi_u(t + t_u)$ in (17) is calculated based on a single MS handover, candidate MSs for offloading should be filtered out by adjusting their ratio values. Suppose a candidate MS u has the target BS b'_u with ABS change t_u, and there are k candidate MSs in total which can be offloaded to that target BS. Then, the adjusted ratio $\bar{\phi}_u(t + t_u)$ can be approximated as

$$\begin{aligned} \bar{\phi}_u(t + t_u) &= \log \frac{G(K_{b'_u} + k) e_{ub'_u}(t + t_u)}{K_{b'_u} + k} \frac{K_{b_u}}{G(K_{b_u}) e_{ub_u}(t)} \\ &= \log \frac{G(K_{b'_u} + 1) e_{ub'_u}(t + t_u)}{K_{b'_u} + 1} \frac{K_{b_u}}{G(K_{b_u}) e_{ub_u}(t)} \\ &\quad + \log \frac{K_{b'_u} + 1}{G(K_{b'_u} + 1)} \frac{G(K_{b'_u} + k)}{K_{b'_u} + k} \\ &\simeq \phi_u(t + t_u) - \frac{k - 1}{K_{b'_u} + 1} \left(1 - \frac{1}{\gamma + \log(K_{b'_u} + 1)} \right), \end{aligned} \quad (19)$$

where b_u is the serving BS of MS u. For proof, please refer to Proposition 1 and 2. When there is only one candidate MS for a target BS (i.e., $k = 1$), $\phi_u(t + t_u)$ is equivalent to $\bar{\phi}_u(t + t_u)$. As k increases, $\bar{\phi}_u(t + t_u)$ decreases accordingly. To maximize the MS offloading gain, we only consider MSs satisfying $\bar{\phi}_u(t + t_u) > 0$ to determine n_i's and m_j's in (10), (12).

n_i's and m_j's can be obtained as follows. In the ABS increment case, for a given target PBS b' and ABS change $+1$, MMSs are sorted in a decreasing order of $\phi_u(t + 1)$. From the first row of the list (i.e., $k = 1$) to the bottom, k is increased by 1 for each row and it is checked if the following condition for the MMS in kth row is met:

$$\phi_u(t + 1) > \frac{k - 1}{K_{b'} + 1} \left(1 - \frac{1}{\gamma + \log(K_{b'} + 1)} \right). \quad (20)$$

Suppose the MMS in the k'th row doesn't satisfy the condition in (20), then the number of MMSs offloaded to PBS b', $m_{b'}$, becomes $k' - 1$. After determining m_j's for all PBSs using this process, n_i's can be determined by checking the serving MBSs of those MMSs. In the ABS decrement case, the target BS type for offloading is macro, therefore we find n_i's first and then m_j's can be determined accordingly.

Upon determining n_i's and m_j's, using the conditions in (10), (12), the number of ABSs can be changed if the following conditions are met:

TABLE II
SIMULATION PARAMETERS

Simulation Parameter	*Value*
Carrier frequency	2.0 GHz
System bandwidth	10 MHz
Antenna configuration	SISO
Channel model	Typical Urban (TU)
Inter-site distance	750 m
Noise power spectral density	-174 dBm/Hz
ABS periodicity (T)	40 subframes
Macro BS transmit power	40 W (46 dBm)
Macrocell path loss model	128.1 + 37.6$\log_{10} R$ (R in km)
Macrocell shadowing model	Log normal fading with std. 10 dB
Macro BS antenna gain	15 dBi
Pico BS transmit power	1 W (30 dBm)
Picocell path loss model	140.7 + 36.7$\log_{10} R$ (R in km)
Picocell shadowing model	Log normal fading with std. 6 dB
Pico BS antenna gain	5 dBi
Min. distance MBS-PBS / PBS-PBS	75 m / 50 m
CRE bias	16 dB
Number of MSs per macrocell	50
Number of picocells per macrocell	2 / 4
Number of simulations	500 times per scenario

- ABS increment by +1: if $\Delta U_{+}^{ABS}(t+1) > \Delta U_{-}^{ABS}(t-1)$ and $\Delta U_{+}^{ABS}(t+1) > \delta_h^{ABS}$
- ABS decrement by -1: if $\Delta U_{-}^{ABS}(t-1) > \Delta U_{+}^{ABS}(t+1)$ and $\Delta U_{-}^{ABS}(t-1) > \delta_h^{ABS}$,

where δ_h^{ABS} is a hysteresis margin to prevent possible ping-pong effects in ABS control.

It should be noted that the proposed method of determining n_i's and m_j's is not optimal for estimating the actual number of MSs that can be offloaded with ABS change. However, from the view point of required report messages from MSs and computations at BSs, the proposed method provides a simple and dynamic way to control the number of ABSs.

V. PERFORMANCE EVALUATION

In this section, we demonstrate the performance of the proposed scheme through simulations. The heterogeneous network deployment is constructed as follows. The macro-tier consists of 7 three-sectorized MBSs (i.e., 21 macrocells), and in each macrocell's coverage outdoor omni-directional picocells and MSs are uniformly distributed. To obtain two achievable link rates c_{ub} and $\bar{c}_{ub}$, we calculate the bit rates based on channel quality indicators (CQIs) fed back from MSs in the system level simulator developed based on the LTE downlink system level simulator in [10]. The detailed parameters are described in Table II, most of which are adopted from 3GPP's system level simulation parameters in [11], [12], [13].

For the performance evaluation, the following schemes are compared through simulations:

TABLE III
AVERAGE NUMBER OF OPTIMAL ABSS

No. of picos	Proposed	CRE	RSS
2 picos	19.1	2.2	0
4 picos	31.9	5.1	0

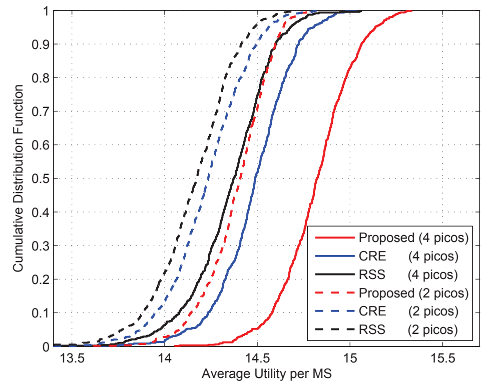

Fig. 1. CDFs of average utility per MS

- *Proposed scheme*: Load-aware cell association and ABS control are jointly optimized.
- *Received signal strength-based cell association (RSS)*: The cell association is done by choosing the cell with the strongest received signal strength.
- *Cell range expansion-based cell association (CRE)*: The cell association is done by choosing the cell with the strongest received signal strength plus the CRE bias.

For the proposed scheme, the initial cell association and ABS duration are given as the received signal strength-based BS selection and zero ABSs, respectively. The hysteresis margins for the load-balancing (stage 1) and the ABS control (stage 2) are set to be $\log 1.1$ and $\log(p \cdot 1.3)$ where p is the number of picocells per macrocell, respectively. For RSS and CRE cases, the optimal number of ABSs is found through exhaustive search that maximizes the network-wide utility.

Table III shows how many ABSs are required for three schemes to maximize the network-wide utility. Due to the large number of MMSs, the RSS scheme achieves the maximum utility with zero ABSs. For the proposed and CRE schemes, the 4 picocell case requires about 50% more ABSs than the 2 picocell case does.

In Fig. 1, the CDF of average utility per MS is shown, which is calculated by dividing the sum of utilities by the total number of MSs in the network. Compared to CRE and RSS schemes, the mean of average utility per MS of the proposed scheme is improved by about 1.5% and 2.5% in 2 picocell and 4 picocell cases, respectively.

Through Table IV and Fig. 2, MS data rate of three schemes are compared. In overall, the proposed scheme shows about +5% $\sim$ +30% and +20% $\sim$ +40% performance improvement in the mean and edge data rate, respectively. In the 2 picocell case, the RSS scheme shows about 5% higher mean data rate

TABLE IV
MS DATA RATE (BPS/HZ)

No. of picos	Metric	Proposed	CRE	RSS
2 picos	Mean	0.226	0.207	0.216
	5%-ile	0.060	0.049	0.043
4 picos	Mean	0.381	0.299	0.293
	5%-ile	0.077	0.054	0.046

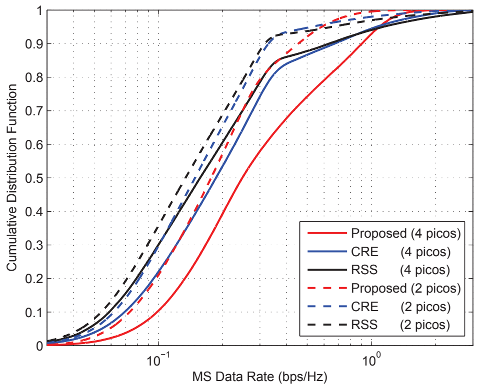

Fig. 2. CDF of MS data rate (bps/Hz)

than the CRE scheme does. This can be explained that the offloading effect in the CRE scheme is limited with 2 picocells so that the higher peak data rate of PMSs in the RSS scheme due to the less number of PMSs per picocell contributes more to the higher mean data rate of the RSS scheme. However, in the 4 picocell case, the CRE scheme shows higher mean data rate than the RSS scheme. Similarly, it is noted in Fig. 2 that above the 95%-ile the proposed scheme shows worse performance than CRE and RSS because the numbers of PMSs in both schemes are much less than the proposed scheme and they can achieve higher data rate.

VI. CONCLUSION

In heterogeneous cellular networks, user offloading between macro- and pico-tier along with the ABS control for cross-tier interference mitigation is important for improving the network capacity. In this paper, the network-wide utility maximization problem has been investigated by jointly optimizing the cell association and the number of ABSs in heterogeneous cellular networks. In the first stage, MS load balancing is done based on their expected data rate by handover under the current number of ABSs t. When there is no MS available for handover, the network-wide net utility by increasing (or decreasing) the number of ABSs is estimated from MSs' expected data rate by handover along with the ABS change in the second stage. Through simulations, we demonstrate that the proposed algorithm not only improves the system performance, but also provides better proportional fairness among users.

ACKNOWLEDGMENT

This material is based upon work partially supported by the National Science Foundation (NSF) grants CNS1035655 and CNS1018346.

APPENDIX

A. Proof of Proposition 1

The net increment of network-wide utility by handing over MS u from BS b to b' can be expressed as

$$\Delta U^{HO} = \left[\log \frac{G(K_{b'}+1)e_{ub'}(t)}{K_{b'}+1} - \log \frac{G(K_b)e_{ub}(t)}{K_b}\right] + \sum_{v \in \mathcal{U}_b \setminus \{u\}} \left[\log \frac{G(K_b-1)e_{vb}(t)}{K_b-1} - \log \frac{G(K_b)e_{vb}(t)}{K_b}\right] + \sum_{w \in \mathcal{U}_{b'}} \left[\log \frac{G(K_{b'}+1)e_{wb'}(t)}{K_{b'}+1} - \log \frac{G(K_{b'})e_{wb'}(t)}{K_{b'}}\right], \quad (21)$$

where the first, second and last term denote the net utility increment of the MS u, the BS b, and b', respectively. Eq. (21) can be further simplified as

$$\Delta U^{HO} = \log \frac{G(K_{b'}+1)e_{ub'}(t)}{K_{b'}+1} \frac{K_b}{G(K_b)e_{ub}(t)} + (K_b-1)\log \frac{G(K_b-1)}{G(K_b)} \frac{K_b}{K_b-1} + K_{b'} \log \frac{G(K_{b'}+1)}{G(K_{b'})} \frac{K_{b'}}{K_{b'}+1}. \quad (22)$$

Using the log function approximation $\log(1 \pm x) \simeq \pm x$ (if x is small and $|x| < 1$) and the Euler's approximation of the multi-user diversity gain $G(k) = \sum_{i=1}^{k} \frac{1}{i} \simeq \gamma + \log k$, where γ $(= 0.5572 \cdots)$ is the Euler-Mascheroni constant, we can obtain the following four equations:

$$\cdot\ (K_b-1)\log \frac{K_b}{K_b-1} = (K_b-1)\log\left(1+\frac{1}{K_b-1}\right) \simeq 1 \quad (23)$$

$$\cdot\ K_{b'} \log \frac{K_{b'}}{K_{b'}+1} = -K_{b'} \log\left(1+\frac{1}{K_{b'}}\right) \simeq -1 \quad (24)$$

$$\begin{aligned}\cdot\ (K_b-1)\log \frac{G(K_b-1)}{G(K_b)} &\simeq (K_b-1)\log \frac{\gamma+\log(K_b-1)}{\gamma+\log K_b} \\ &\simeq (K_b-1)\log\left(1+\frac{\log(K_b-1)-\log K_b}{\gamma+\log K_b}\right) \\ &\simeq (K_b-1)\log\left(1+\frac{\log\left(1-\frac{1}{K_b}\right)}{\gamma+\log K_b}\right) \\ &\simeq (K_b-1)\log\left(1-\frac{\frac{1}{K_b}}{\gamma+\log K_b}\right) \\ &\simeq -\frac{K_b-1}{K_b}\frac{1}{\gamma+\log(K_b)} \simeq \frac{-\left(1-\frac{1}{K_b}\right)}{\gamma+\log K_b} \end{aligned} \quad (25)$$

$$\cdot\, K_{b'} \log \frac{G(K_{b'}+1)}{G(K_{b'})} \simeq K_{b'} \log \frac{\gamma + \log(K_{b'}+1)}{\gamma + \log K_{b'}}$$
$$\simeq K_{b'} \log \left(1 + \frac{\log(K_{b'}+1) - \log K_{b'}}{\gamma + \log K_{b'}}\right)$$
$$\simeq K_{b'} \log \left(1 + \frac{\log\left(1+\frac{1}{K_{b'}}\right)}{\gamma + \log K_{b'}}\right) \simeq \frac{1}{\gamma + \log K_{b'}}. \quad (26)$$

By plugging equations in (23)-(26) into (22), Eq. (22) becomes

$$\Delta U^{HO} = \log \frac{G(K_{b'}+1)e_{ub'}(t)}{K_{b'}+1} \frac{K_b}{G(K_b)e_{ub}(t)} + \frac{1}{\gamma + \log K_{b'}} - \frac{\left(1 - \frac{1}{K_b}\right)}{\gamma + \log K_b}. \quad (27)$$

In order for the user's handover to improve the network-wide utility, $\Delta U^{HO} > 0$ needs to be satisfied, therefore we can derive the condition in (8).

B. Proof of Proposition 2 & 3

For the ABS increment case, $\Delta U_m(t+1)$ and $\Delta U_p(t+1)$ are derived as

$$\Delta U_m(t+1) = \sum_{i\in\mathcal{B}_m} \sum_{w\in\mathcal{U}_i\setminus\mathcal{U}_i^{OL}} \left[\log \frac{G(K_i-n_i)e_{wi}(t+1)}{K_i-n_i} - \log \frac{G(K_i)e_{wi}(t)}{K_i}\right]$$
$$\simeq \sum_{i\in\mathcal{B}_m} \left[-\frac{n_i\left(1-\frac{n_i}{K_i}\right)}{\gamma+\log K_i} + n_i - \frac{K_i - n_i}{T-t}\right], \quad (28)$$

$$\Delta U_p(t+1) = \sum_{j\in\mathcal{B}_p} \sum_{v\in\mathcal{U}_j} \left[\log \frac{G(K_j+m_j)e_{vj}(t+1)}{K_j+m_j} - \log \frac{G(K_j)e_{vj}(t)}{K_j}\right]$$
$$\simeq \sum_{j\in\mathcal{B}_p} \left[\frac{m_j}{\gamma+\log K_j} - m_j + \sum_{v\in\mathcal{U}_j} \frac{a_{vj}-1}{T+t(a_{vj}-1)}\right], \quad (29)$$

respectively, using approximations in (23)-(26) and the followings for $i \in \mathcal{B}_m$ and $j \in \mathcal{B}_p$:

$$\cdot \log \frac{e_{wi}(t+1)}{e_{wi}(t)} = \log \frac{\left(1-\frac{t+1}{T}\right)c_{wi}}{\left(1-\frac{t}{T}\right)c_{wi}} = \log \frac{T-t-1}{T-t}$$
$$= \log\left(1 - \frac{1}{T-t}\right) \simeq -\frac{1}{T-t}, \quad (30)$$

$$\cdot \log \frac{e_{vj}(t+1)}{e_{vj}(t)} = \log \frac{\left(1-\frac{t+1}{T}\right)c_{vj} + \frac{t+1}{T}\bar{c}_{vj}}{\left(1-\frac{t}{T}\right)c_{vj} + \frac{t}{T}\bar{c}_{vj}}$$
$$= \log\left(1 + \frac{\bar{c}_{vj} - c_{vj}}{Tc_{vj} + t(\bar{c}_{vj} - c_{vj})}\right) \simeq \frac{a_{vj}-1}{T+t(a_{vj}-1)} \quad (31)$$

Since $\sum_{i\in\mathcal{B}_m} n_i - \sum_{j\in\mathcal{B}_p} m_j = 0$, the requirement $\Delta U_+^{ABS}(t+1) > 0$ for ABS increment leads us to the condition in (10).

Similarly, $\Delta U_m(t-1)$ and $\Delta U_p(t-1)$ for the ABS decrement case can be approximated as

$$\Delta U_m(t-1) = \sum_{i\in\mathcal{B}_m} \sum_{w\in\mathcal{U}_i} \left[\log \frac{G(K_i+n_i)e_{wi}(t-1)}{K_i+n_i} - \log \frac{G(K_i)e_{wi}(t)}{K_i}\right]$$
$$\simeq \sum_{i\in\mathcal{B}_m} \left[\frac{n_i}{\gamma+\log K_i} - n_i + \frac{K_i}{T-t}\right], \quad (32)$$

$$\Delta U_p(t-1) = \sum_{j\in\mathcal{B}_p} \sum_{v\in\mathcal{U}_j\setminus\mathcal{U}_j^{OL}} \left[\log \frac{G(K_j-m_j)e_{vj}(t-1)}{K_j-m_j} - \log \frac{G(K_j)e_{vj}(t)}{K_j}\right]$$
$$\simeq \sum_{j\in\mathcal{B}_p} \left[-\frac{m_j\left(1-\frac{m_j}{K_j}\right)}{\gamma+\log K_j} + m_j - \sum_{v\in\mathcal{U}_j\setminus\mathcal{U}_j^{OL}} \frac{a_{vj}-1}{T+t(a_{vj}-1)}\right], \quad (33)$$

respectively.

References

[1] S. Landstrom, A. Furuskar, K. Johansson, L. Falconetti, and F. Kronestedt, "Heterogeneous networks - increasing cellular capacity," *Ericson Review*, vol. 89, pp. 4–9, 2011.

[2] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3gpp heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 10–21, June 2011.

[3] T. Bu, L. Li, and R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, April 2006, pp. 1–12.

[4] K. Son, S. Chong, and G. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 7, pp. 3566–3576, July 2009.

[5] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. Andrews, "User association for load balancing in heterogeneous cellular networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 6, pp. 2706–2716, June 2013.

[6] S. Corroy, L. Falconetti, and R. Mathar, "Dynamic cell association for downlink sum rate maximization in multi-cell heterogeneous networks," in *Communications (ICC), 2012 IEEE International Conference on*, June 2012, pp. 2457–2461.

[7] Q. Ye, M. Al-Shalash, C. Caramanis, and J. Andrews, "On/off macrocells and load balancing in heterogeneous cellular networks," in *Global Telecommunications Conference (GLOBECOM 2013), 2013 IEEE*, Dec. 2013, pp. 3919–3924.

[8] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, 1997.

[9] H. Kushner and P. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 4, pp. 1250–1259, July 2004.

[10] J. C. Ikuno, M. Wrulich, and M. Rupp, "System level simulation of LTE networks," in *Proc. 2010 IEEE 71st Vehicular Technology Conference*, Taipei, Taiwan, May 2010.

[11] 3rd Generation Partnership Project, "3GPP TR 36.912 V11.0.0 - Feasibility study for Further Advancements for E-UTRA (LTE-Advanced) (Release 11)," Tech. Rep., Sep. 2012.

[12] ——, "3GPP TR 36.814 V9.0.0 - Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects," Tech. Rep., Mar. 2010.

[13] ——, "3GPP TR 36.931 V11.0.0 - Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) requirements for LTE Pico Node B," Tech. Rep., Sep. 2012.