

Enhancing Performance of Heterogeneous Cloud Radio Access Networks with Efficient User Association

Shaowei Wang and Yang Sun

School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China

E-mail: wangsw@nju.edu.cn, flxzaj@163.com

Abstract—Heterogeneous cloud radio access network (H-CRAN) is proposed as a cost-effective paradigm to meet the ever-increasing mobile data traffic demand, where the key idea is applying cloud computing technologies in a heterogeneous network (HetNet) to improve both spectral and energy efficiencies of the cellular system. In this paper, we investigate how to provide as many as possible users with QoS-guaranteed services in the H-CRAN. Our optimization task is to maximize the number of users with rate requirements for a given set of access points with bandwidth and transmission power budgets. We develop a novel user association strategy, where we present a *Reference Power* concept and develop an approximation algorithm to address the intractable user association problem. Numerical results indicate that our proposed user association strategy can not only increase the fully satisfied users significantly as compared to other methods, but also fulfill the capacity potential of the H-CRAN.

I. INTRODUCTION

ONE of the most critical challenges facing fifth generation (5G) mobile communication network is that of satisfying growing traffic demands, which receives great attention recently [1]. Heterogeneous network (HetNet) is a promising scheme to achieve this goal [2], where densely deployed small cells coexist with traditional macro base stations (BSs) to improve the system throughput by exploiting spatial spectrum reuse gain. However, co-tier and cross-tier interference management becomes a big challenge for the HetNet. Moreover, deploying massive BSs raise new issues about network planning and optimization [3], including the burden of capital expenditure and operating expense,

Cloud radio access network (C-RAN) is introduced as another promising architecture that has the potential to overcome these challenges [4]. In the C-RAN, baseband processing is separated from remote radio heads (RRHs) and centralized in a baseband unit (BBU) pool. The BBU pool creates a set of common processing resources that can enable efficient resource allocation and interference managements and increase the flexibility for network upgrade. However, it requires high rate backhauls between RRHs and the BBU pool, which impedes the scalability of the C-RAN. Moreover, there exists a potential coverage issue for the C-RAN since the RRHs

usually transmit with much smaller power as compared to the macro BSs in the HetNet or the conventional cellular system.

To overcome the challenges facing HetNet and C-RAN, heterogeneous cloud radio access network (H-CRAN) is proposed as an alternative infrastructure for future mobile networks [5, 6]. By incorporating cloud computing into the HetNet, the H-CRAN can reduce inter-tier interference and improve cooperative processing gains, which makes full use of the advantages of both Hetnet and C-RAN. An H-CRAN usually consists of multiple macro BSs and a large number of RRHs which are connected with the BBU pool via backhaul links. With cooperative processing techniques based on cloud computing, macro BSs can cooperate with the BBU pool to weaken the cross-tier interference between RRHs and macro BSs. Compared to the C-RAN, part of the control and broadcast functionalities in the H-CRAN can be shifted from the RRHs/BBU pool to the macro BSs, which alleviates the burdens of the backhauls between RRHs and the BBU pool. Moreover, macro BSs can provide seamless coverage and support services for the users moving at high speed or requesting low transmission rates. Meanwhile, RRHs are densely deployed in hot spots to provide high transmission rates for users with high quality of service (QoS) requirements. Furthermore, the macro BSs in the H-CRAN can guarantee the backward compatibility with the existing cellular networks, making it more practical to upgrade from current 2G/3G/4G cellular systems.

A nonnegligible issue arises from the H-CRAN naturally: When a user requests to access the network, which access point (macro BS or RRH) should be assigned to serve the user? In other words, user association should be carefully designed to fulfill the potential of the H-CRAN. Many works on user association in cellular radio networks have been proposed in the literature. Maximum received signal power (MRSP) or signal-to-interference-plus-noise ratio (SINR) based schemes are discussed extensively [7–9], where the user always associated with the access point providing the largest MRSP/SINR. Cell range expansion (CRE) is a promising offloading technology for the HetNet [10, 11], where a positive range expansion bias is added to the received signals of users to increase coverage footprints of low-power access nodes, e.g., the RRHs. However, due to the variation of spatial traffic, the value of bias should be changed dynamically to provide acceptable system performance.

In this paper, we propose an effective user association

This work was partially supported by the National Natural Science Foundation of China (61671233), the Jiangsu Science Foundation (BK20151389) and the open research fund of National Mobile Communications Research Laboratory (2016D08).

strategy for the H-CRAN, where rate requirements of users, power and bandwidth constraints of macro BSs and RRHs are taken into consideration. Our optimization objective is to maximize the number of fully satisfied users with strained radio resources. The main contributions of this work are summarized as follows:

- We study a basic bandwidth and power allocation problem, minimizing the required power to satisfy the rate requirements of a given set of users for each RRH. A fast algorithm is developed by exploiting the special structure of Hessian matrix. The storage complexity and computational complexity can be substantially reduced compared to standard algorithms.
- We design a $\frac{1}{2}$ -approximation algorithm for the user association problem by introducing the *Reference Power* concept. Different from traditional MRSP/SINR-based and CRE schemes, the user with the minimum *Reference Power* has the priority to access the corresponding RRH. Numerical results validate that our proposed user association strategy can increase the fully satisfied users compared with other schemes.

The remainder of this paper is organized as follows. In Section II, system model and problem formulation are illustrated. In Section III, our proposed algorithms are presented in detail. In Section IV, numerical results are given with discussions. Conclusions are drawn in Section V.

II. NETWORK MODEL AND PROBLEM FORMULATION

A. System Model

To make the rest of this paper easy to follow, we list frequently used notations in Table I.

Consider a region $\mathcal{D} \in \mathbf{R}^2$ served by an H-CRAN including k users, a macro BS and $(N-1)$ RRHs associated with a BBU pool. The set of access points, including RRHs and the macro BS, is denoted by $\mathcal{N} = \{1, 2, \dots, N\}$, where the last element $N \in \mathcal{N}$ represents the macro BS. For access point $n \in \mathcal{N}$, its maximum transmission power is p_n^{max} and the total available bandwidth is b_n^{max} . To simplify discussions, we assume that the macro BS and the RRHs differ mainly in transmission power.

The set of users is denoted by $\mathcal{K} = \{1, 2, \dots, K\}$. The minimal rate requirement of user $k \in \mathcal{K}$ is R_k^{min} . Denote $\rho_{k,n}$ as the assignment index indicating whether user k is served by RRH n or not:

$$\rho_{k,n} = \begin{cases} 1 & \text{user } k \text{ is served by RRH } n, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Denote $h_{k,n}$ as the power gain between the RRH n and the user k . Let $b_{k,n}$ and $p_{k,n}$ be the bandwidth and power of the RRH n allocated to the user k , respectively. If the user k is served by the RRH n , i.e. $\rho_{k,n} = 1$, the maximum interference introduced by adjacent active RRHs with unit bandwidth is

$$I_{k,n} = \sum_{n' \in \mathcal{N}, n' \neq n} p_{n'}^{max} h_{k,n'} / b_{n'}^{max}, \quad (2)$$

TABLE I
NOTATIONS

Symbol	Semantics
$b_{k,n}$	Bandwidth of RRH n allocated to user k
b_n^{max}	Total available bandwidth of RRH n
$h_{k,n}$	Power gain between RRH n and user k
\mathcal{K}	Set of users
K	Number of users
$I_{k,n}$	Maximum interference introduced by adjacent RRHs with unit bandwidth to user k served by RRH n
\mathcal{N}	Set of RRHs
N	Number of RRHs
N_0	PSD of AWGN
$p_{k,n}$	Power of RRH n allocated to user k
p_n^{max}	Maximum transmission power of RRH n
$r_{k,n}$	Lower bound of achievable rate from RRH n to user k
R_k^{min}	Rate requirement of user k
$\rho_{k,n}$	User association index between user k and RRH n
Φ_k	Selection variable of user k

where $\mathcal{N}_s = \{n \in \mathcal{N} | z_n = 1\}$ is the set of active RRHs. Thus, the lower bound of achievable rate between the RRH n and the user k can be calculated as

$$r_{k,n} = b_{k,n} \log_2 \left[1 + \frac{p_{k,n} h_{k,n}}{b_{k,n} (N_0 + I_{k,n})} \right], \quad (3)$$

where N_0 is the power spectral density (PSD) of additive white Gaussian noise (AWGN).

For notation brevity, we collect variables $b_{k,n}$'s, $p_{k,n}$'s and $r_{k,n}$'s into vectors \bar{b} , \bar{p} and \bar{r} .

B. Problem Formulation

It is intuitive that the RRHs should serve as more as possible users with QoS guarantee so that the number of required RRHs could be minimized to save power.

The optimization objective of user association is to maximize the number of satisfied users for a given set of RRHs, where each user can be associated with at most one RRH. We use Φ_k to indicate whether user k can be served by the H-CRAN or not, that is,

$$\Phi_k = \sum_{n \in \mathcal{N}} \rho_{k,n} = \begin{cases} 1 & \text{user } k \text{ served by the H-CRAN,} \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Given a set of RRHs \mathcal{N} , the user association problem is mathematically formulated as follows:

$$\begin{aligned} & \text{maximize} && \sum_{k \in \mathcal{K}} \Phi_k \\ & \text{s.t. } C_1 : && \sum_{k \in \mathcal{K}} b_{k,n} \leq b_n^{max}, \forall n \in \mathcal{N}, \\ & C_2 : && \sum_{k \in \mathcal{K}} p_{k,n} \leq p_n^{max}, \forall n \in \mathcal{N}, \\ & C_3 : && \sum_{n \in \mathcal{N}} \rho_{k,n} r_{k,n} = \Phi_k R_k^{min}, \forall k \in \mathcal{K}, \\ & C_4 : && b_{k,n} \geq 0, p_{k,n} \geq 0, \forall k \in \mathcal{K}, \forall n \in \mathcal{N}, \\ & C_5 : && \Phi_k, \rho_{k,n} \in \{0, 1\}, \forall k \in \mathcal{K}, n \in \mathcal{N}. \end{aligned} \quad (5)$$

C_1 and C_2 are the bandwidth and transmission power budgets of the RRHs. C_3 ensures that the RRHs should satisfy user's

rate requirement if $\Phi_k = 1$. Obviously, if R_k^{min} is the same for all users, problem (5) is equivalent to maximizing system capacity that has been extensively investigated in the literature.

III. OUR PROPOSED ALGORITHMS

A. Fast Algorithm for Bandwidth and Power Allocation

First, we need to answer a question: Is it possible for the RRH n to satisfy a given set of users \mathcal{K}_n with the bandwidth and power budgets of the RRH n . This feasibility problem can be equivalent to the following task: Given available bandwidth b_n^{max} , how to work out the minimum power consumption of the RRH n while satisfying the rate requirements of users in set \mathcal{K}_n . Such a problem can be written as follows:

$$\begin{aligned} \min_{b_{k,n}, p_{k,n}} \quad & \sum_{k \in \mathcal{K}_n} p_{k,n} \\ \text{s.t. } C_1: \quad & \sum_{k \in \mathcal{K}_n} b_{k,n} = b_n^{max}, \\ C_2: \quad & r_{k,n} = R_k^{min}, \forall k \in \mathcal{K}_n, \\ C_3: \quad & b_{k,n} \geq 0, p_{k,n} \geq 0, \forall k \in \mathcal{K}_n. \end{aligned} \quad (6)$$

If the optimal value of the problem does not exceed $p_{n,max}$, all rate requirements of the users in \mathcal{K}_n can be satisfied by RRH n . According to the constraints in C_2 of (6) and the definition of $r_{k,n}$, we can obtain

$$p_{k,n} = \frac{b_{k,n}}{H_{k,n}} \cdot \left(2^{R_k^{min}/b_{k,n}} - 1 \right), \quad (7)$$

where $H_{k,n} = h_{k,n}/(N_0 + I_{k,n})$. Note that $I_{k,n}$ is fixed in problem (6). Substituting (7) into problem (6), we have

$$\begin{aligned} \min_{b_{k,n}} \quad & \sum_{k \in \mathcal{K}_n} \frac{b_{k,n}}{H_{k,n}} \cdot \left(2^{R_k^{min}/b_{k,n}} - 1 \right) \\ \text{s.t. } C_1: \quad & \sum_{k \in \mathcal{K}_n} b_{k,n} \leq b_n^{max}, \\ C_2: \quad & b_{k,n} \geq 0, \forall k \in \mathcal{K}_n. \end{aligned} \quad (8)$$

It is easy to prove that (6) defines a convex problem because the objective function is convex and all the constraints are affine [12]. It can be solved by standard convex optimization techniques, such as barrier method [12] that is quadratic convergence. However, both the storing cost and the computing cost in the barrier method are high. For problem (8), its storage complexity and computation complexity yielded by standard barrier method are $O(|\mathcal{K}_n|^2)$ and $O(|\mathcal{K}_n|^3)$, respectively. Obviously, it is undesirable. We design a fast algorithm to substantially reduce the storage and computation complexity.

Collect all variables into one vector $\bar{b}_n \in \mathbf{R}^{|\mathcal{K}_n| \times 1}$ and define

$$b_n = b_n^{max} - \sum_{k \in \mathcal{K}_n} b_{k,n}, \quad (9)$$

we can convert all inequality constraints into a logarithmic barrier function $\phi(\bar{b}_n)$:

$$\phi(\bar{b}_n) = -\log b_n - \sum_{k \in \mathcal{K}_n} \log b_{k,n}. \quad (10)$$

Problem (8) can be converted into a sequence of minimization problems by introducing a logarithmic barrier function with a

parameter t . The optimal solution to (8) can be approximated by solving the following unconstrained convex problem:

$$\min_{\bar{b}_n} \psi_t(\bar{b}_n) = t \sum_{k \in \mathcal{K}_n} p_{k,n} + \phi(\bar{b}_n) \quad (11)$$

Generally, Newton method is preferred to solve the convex problem with equality constraints due to its quadratic convergence property [12]. For a given parameter t , Newton step $\Delta \bar{b}_n$ can be calculated by solving the following equation:

$$\nabla^2 \psi_t(\bar{b}_n) \Delta \bar{b}_n = -\nabla \psi_t(\bar{b}_n). \quad (12)$$

where $\nabla^2 \psi_t(\bar{b}_n)$ and $\nabla \psi_t(\bar{b}_n)$ are the Hessian and the gradient of $\psi_t(\bar{b}_n)$, respectively. The Hessian matrix can be written as

$$\nabla^2 \psi_t(\bar{b}_n) = \Pi + \frac{\nabla b_n \nabla b_n^T}{b_n^2}, \quad (13)$$

where $\Pi = \text{diag}(\Pi_1, \Pi_2, \dots, \Pi_{|\mathcal{K}_n|})$ with

$$\Pi_k = t \cdot \frac{\partial^2 p_{k,n}}{\partial b_{k,n}^2} + \frac{1}{b_{k,n}^2}. \quad (14)$$

Fact 1. Given a non-singular matrix $\Pi \in \mathbf{R}^{|\mathcal{K}_n| \times |\mathcal{K}_n|}$, vectors $f, b \in \mathbf{R}^{|\mathcal{K}_n| \times 1}$, where f satisfies $1 + f^T \Pi^{-1} f \neq 0$. Then, if $\Pi x = b$, $(\Pi + f f^T) \tilde{x} = b$, it always holds $\tilde{x} = x - \frac{\langle f, x \rangle}{1 + \langle f, g \rangle} g$, where $g = \Pi^{-1} f$, $g \in \mathbf{R}^{|\mathcal{K}_n| \times 1}$.

By using the mathematical fact, we can calculate Newton step efficiently. First, we can obtain \bar{x}_1 and \bar{x}_2 by solving $\Pi \bar{x}_1 = -\nabla \psi_t(\bar{x}_n)$ and $\Pi \bar{x}_2 = \bar{\xi}$, where $\bar{\xi} = \nabla b_n / b_n$. Since D is a diagonal matrix, \bar{x}_1 and \bar{x}_2 can be solved with complexity of $O(|\mathcal{K}_n|)$. Then, Newton step $\Delta \bar{b}_n$ can be calculated as

$$\Delta \bar{b}_n = \bar{x}_1 - \frac{\langle \bar{\xi}, \bar{x}_1 \rangle}{1 + \langle \bar{\xi}, \bar{x}_2 \rangle} \bar{x}_2. \quad (15)$$

The computation complexity of our proposed method is $O(L|\mathcal{K}_n|)$, where L is the number of Newton iterations. Moreover, the storage complexity is bounded by $O(|\mathcal{K}_n|)$ since only diagonal elements of Hessian, one vector and gradient need to be stored. The outline of our proposed algorithm is summarized in Table II. ϵ_b and ϵ_n are the tolerances of barrier method and Newton method, respectively. α and β are two constants utilized in backtracking line search with $\alpha \in (0, 0.5)$ and $\beta \in (0, 1)$. The step size of backtracking line search is s with $s > 0$. t and μ are parameters associated with a trade-off between outer iterations and inner iterations.

Define $p_n(\mathcal{K}_n) = \sum_{k \in \mathcal{K}_n} p_{k,n}^*$ be the optimal value of (8). If $p_n(\mathcal{K}_n)$ does not exceed $p_{n,max}$, all rate requirements of the users in \mathcal{K}_n can be satisfied by RRH n ; otherwise, RRH n cannot serve all users in \mathcal{K}_n .

B. A $\frac{1}{2}$ -Approximation Algorithm for User Association

Define $p_n(\{k\})$ as Reference Power of the user k associated with the RRH n :

$$p_n(\{k\}) = \frac{b_n^{max}}{H_{k,n}} \cdot \left(2^{R_k^{min}/b_n^{max}} - 1 \right). \quad (16)$$

The Reference Power $p_n(\{k\})$ indicates the minimum power to satisfy the rate requirement of user k for RRH n .

TABLE II
BARRIER METHOD

Algorithm 1	
1:	<i>Initialization:</i> Feasible point $\bar{b}_n, t, \epsilon_b, \epsilon_n, \alpha, \beta, \mu$;
2:	while $ \mathcal{K}_n /t > \epsilon_b$
3:	while $\lambda^2/2 > \epsilon_n$
4:	Obtain $\Delta\bar{b}_n$ and set $s = 1$;
5:	while $\psi_t(\bar{b}_n + s\Delta\bar{b}_n) > \psi_t(\bar{b}_n) - \alpha s\lambda^2$
6:	$s = \beta s$;
7:	end while
8:	Update $\bar{b}_n = \bar{b}_n + s\Delta\bar{b}_n$;
9:	end while
10:	$t = \mu t$;
11:	end while
12:	return \bar{b}_n

TABLE III
 $\frac{1}{2}$ -APPROXIMATION ALGORITHM FOR USER ASSOCIATION PROBLEM

Algorithm 2	
1:	<i>Initialization:</i> $\mathcal{K}_n = \emptyset, \forall n \in \mathcal{N}; \mathcal{K}_{temp} = \mathcal{K}; \mathcal{N}_{temp} = \mathcal{N}$;
2:	Calculate the reference power $p_n(\{k\}), k \in \mathcal{K}, n \in \mathcal{N}$;
3:	repeat
4:	$(k', n') = \arg \min_{(k,n): k \in \mathcal{K}_{temp}, n \in \mathcal{N}_{temp}} P_n(\{k\})$;
5:	if $P_{n'}(\mathcal{K}_{n'} \cup \{k'\}) \leq p_{n'}^{max}$
6:	$\mathcal{K}_{n'} \leftarrow \mathcal{K}_{n'} \cup \{k'\}$;
7:	$\mathcal{K}_{temp} \leftarrow \mathcal{K}_{temp} \setminus \{k'\}$;
8:	else
9:	$\mathcal{N}_{temp} \leftarrow \mathcal{N}_{temp} \setminus \{n'\}$;
10:	end if
12:	until $\mathcal{K}_{temp} = \emptyset$ or $\mathcal{N}_{temp} = \emptyset$
13:	return \mathcal{K}_n

We develop an approximation algorithm to obtain promising solutions to the user association problem. First, we initialize $\mathcal{K}_n = \emptyset$ for each RRH and calculate $p_n(\{k\})$ for each $k \in \mathcal{K}, n \in \mathcal{N}$. Then, we find out (k', n') that corresponds to the minimum power cost $p_{n'}(\{k'\})$ and calculate $p_{n'}(\mathcal{K}_{n'} \cup \{k'\})$ by using the fast barrier method. If $p_{n'}(\mathcal{K}_{n'} \cup \{k'\}) \leq p_{n'}^{max}$, we assign user k' to RRH n' , i.e. $\mathcal{K}_{n'} = \mathcal{K}_{n'} \cup \{k'\}$; otherwise, RRH n' cannot serve any remaining user since user k' can use the same allocated bandwidth with less power consumption compared to any other user. Such a procedure terminates when all users have been allocated or all RRHs cannot serve any other user. The algorithm is described in Table III.

Lemma 1. *Given sets of users, denoted by $\mathcal{M}_1, \mathcal{M}_2$, where $|\mathcal{M}_1| = |\mathcal{M}_2|$. For RRH n , if $p_n(\{k_1\}) \geq p_n(\{k_2\}), \forall k_1 \in \mathcal{M}_1, k_2 \in \mathcal{M}_2$, we have $p_n(\mathcal{M}_1) \geq p_n(\mathcal{M}_2)$.*

Proof: The proof is presented in Appendix A. ■

It is straightforward to obtain the following corollary: Given RRH n and sets of users $\mathcal{M}_1, \mathcal{M}_2$. If $p_n(\{k_1\}) \geq p_n(\{k_2\}), \forall k_1 \in \mathcal{M}_1, k_2 \in \mathcal{M}_2$, and $p_n(\mathcal{M}_1) < p_n(\mathcal{M}_2)$, then $|\mathcal{M}_1| < |\mathcal{M}_2|$ holds. The conclusion is intuitive because if $|\mathcal{M}_1| \geq |\mathcal{M}_2|$, we can always find a subset $\mathcal{M}' \subseteq \mathcal{M}_1$ of users that satisfies $p_n(\mathcal{M}_1) \geq p_n(\mathcal{M}') \geq p_n(\mathcal{M}_2)$ based on Lemma 1, where $|\mathcal{M}'| = |\mathcal{M}_2|$.

Theorem 1. *Algorithm 2 is $\frac{1}{2}$ -approximation for the user association problem.*

Proof: Define \mathcal{K}^* as the set of fully satisfied users corresponding to the optimal solution. For each RRH $n \in \mathcal{N}$, define \mathcal{K}_n^* as the set of users satisfied by RRH n in the optimal solution.

Let \mathcal{K}' be the set of satisfied users chosen by Algorithm 2 and \mathcal{K}'_n be the set of users assigned to RRH $n \in \mathcal{N}$. According to the user association procedure shown in Algorithm 2, the user assigned to the RRH n has less power consumption compared to the user in $\mathcal{K}_n^* \setminus \mathcal{K}'$. We have

$$p_n(\{k_2\}) \geq p_n(\{k_1\}), \forall k_1 \in \mathcal{K}'_n, k_2 \in \mathcal{K}_n^* \setminus \mathcal{K}'. \quad (17)$$

Moreover, we can obtain

$$p_n(\mathcal{K}'_n \cup \{k_2\}) > p_n^{max}, \quad (18)$$

for each user $k_2 \in \mathcal{K}_n^* \setminus \mathcal{K}'$ since the user k_2 would be assigned to the RRH n otherwise. Since the users in \mathcal{K}_n^* are associated with the RRH n , we also have

$$p_n^{max} \geq p_n(\mathcal{K}_n^* \setminus \mathcal{K}'). \quad (19)$$

Combining (18) with (19), we can obtain

$$p_n(\mathcal{K}'_n \cup \{k_2\}) > p_n(\mathcal{K}_n^* \setminus \mathcal{K}'). \quad (20)$$

Based on (17), (20) and the corollary discussed above, we have

$$|\mathcal{K}_n^* \setminus \mathcal{K}'| < |\mathcal{K}'_n \cup \{k_2\}| = |\mathcal{K}'_n| + 1, \quad (21)$$

which is equivalent to

$$|\mathcal{K}'_n| \geq |\mathcal{K}_n^* \setminus \mathcal{K}'|. \quad (22)$$

Then,

$$\begin{aligned} 2 \cdot |\mathcal{K}'| &= |\mathcal{K}'| + \sum_{n \in \mathcal{N}} |\mathcal{K}'_n| \\ &\geq |\mathcal{K}^* \cap \mathcal{K}'| + \sum_{n \in \mathcal{N}} |\mathcal{K}_n^* \setminus \mathcal{K}'| \\ &= |\mathcal{K}^* \cap \mathcal{K}'| + |\mathcal{K}^* \setminus \mathcal{K}'|. \end{aligned} \quad (23)$$

Note that $\mathcal{K}^* \cap \mathcal{K}'$ denotes the selected users in \mathcal{K}^* and $\mathcal{K}^* \setminus \mathcal{K}'$ denotes the unselected users in \mathcal{K}^* . Thus $(\mathcal{K}^* \cap \mathcal{K}') \cup (\mathcal{K}^* \setminus \mathcal{K}') = \mathcal{K}^*$, i.e. $|\mathcal{K}^* \cap \mathcal{K}'| + |\mathcal{K}^* \setminus \mathcal{K}'| = |\mathcal{K}^*|$, and finally we get

$$|\mathcal{K}'| \geq \frac{1}{2} |\mathcal{K}^*|. \quad (24)$$

In summary, our proposed user association strategy follows such an idea: User k is associated with the RRH n with the minimum *reference power* that is related to three system parameters: available bandwidth of RRH n , the rate requirement of user k and the SINR with unit power.

IV. NUMERICAL RESULTS

In this section, we give numerical results to evaluate the performance of our proposed algorithms. Simulation parameters of our considered H-CRAN, such as path-loss model, maximum transmission power, system bandwidth, spectral efficiency requirement, etc., are based on the specifications proposed in [13]. All results are averaged over 200 Monte Carlo simulations. In the H-CRAN, the RRHs are distributed uniformly in the region with area of 2×2 km². The maximum

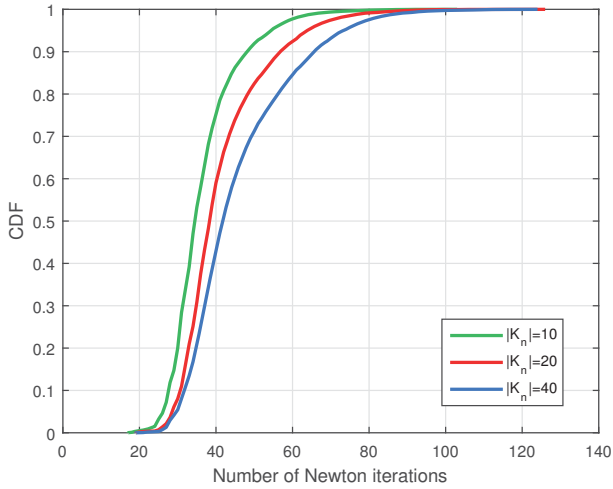


Fig. 1. CDF of the number of Newton iterations required for convergence. $t^{(0)} = 0.1, \epsilon_b = \epsilon_n = 10^{-3}, \mu = 10, \alpha = 0.01, \beta = 0.1$.

transmission power of the macro BS is 40W while the transmission power of each RRH is limited to 1W. According to the technical report in [14], we do not consider the cross-tier interference between macro BS and RRHs. The bandwidth of each RRH is randomly chosen from [20 40 60 80 100] MHz. The macro BS is always active in order to provide basic wireless coverage. The users are also distributed uniformly in the region. The rate requirement of each user is randomly selected from [0.3 3 30] Mbps. Path loss (in dB) from macro BS to user is calculated as $128.1 + 37.6 \log_{10}(D)$. Path loss from RRHs to user is calculated as $140.7 + 36.7 \log_{10}(D)$. Here D (in km) is the distance between user and access point. The standard deviation of lognormal shadowing is 10 dB. The noise PSD is -184 dBm/Hz.

First, we study the convergence of our proposed fast barrier method. Fig. 1 shows the cumulative distribution function (CDF) of Newton iterations. Parameters of the barrier method are set as follows: $t^{(0)} = 0.1, \epsilon_b = \epsilon_n = 10^{-2}, \mu = 10, \alpha = 0.01, \beta = 0.1$. As can be seen from Fig. 1, 95% of Newton iterations are less than 80 for all cases. Moreover, the number of Newton iterations only increases slightly as increasing of the number of users. We can conclude that the fast barrier method converge rapidly and stably.

Second, we evaluate the performance of our proposed user association strategy defined by Algorithm 2. Three user association schemes are considered for comparison: SINR-based user association; CRE with bias of 6dB; CRE with bias of 12dB [10]. For the SINR-based user association strategy, user k is assigned to an RRH or the macro BS that can provide the maximum SINR. For the considered system, user k will be assigned to $n^* = \arg \max_{n \in \mathcal{N}_s} p_n^{max} H_{k,n}$. For the CRE scheme, a positive range expansion bias, 6dB or 12dB, is added to the received signals at users served by the RRHs.

The percentage of satisfied users (i.e. $\sum_k \Phi_k / K$) with different number of users is shown Fig. 2, where the number of access points is $N = 20$. As can be seen in Fig. 2, the SINR-based user association performs poorly in the H-CRAN.

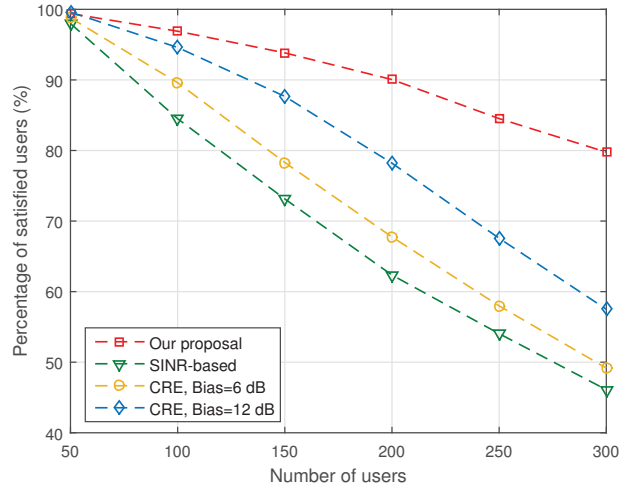


Fig. 2. Percentage of satisfied users as a function of number of users. $N = 20$.

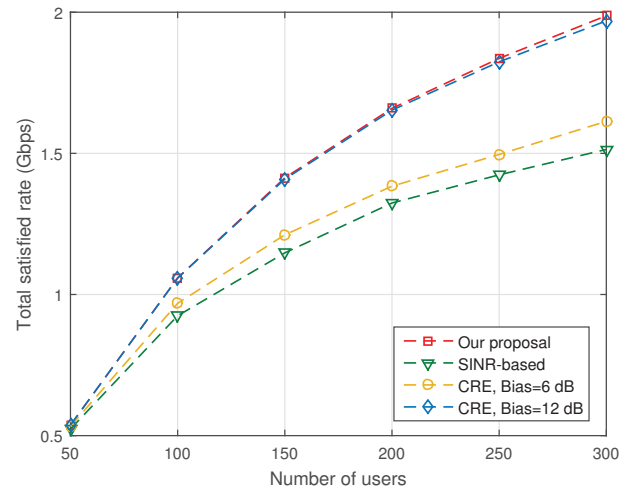


Fig. 3. Satisfied rate requirements as a function of the number of users. $N = 20$.

It is intuitive because most of the users are associated with the macro BS, resulting that RRHs can only offload a small number of users. In contrary, more users can be offloaded to the RRHs by the CRE schemes. From Fig. 2, we can see our proposed user association strategy outperforms the CRE schemes. The performance gap enlarges as the increasing of users. Specifically, more than 80% users can successfully access the H-CRAN by using our proposal when $K = 300$. We can conclude that the proposed user association strategy works well even though radio resources are strained.

The satisfied rate (i.e. $\sum_k R_k^{min} \Phi_k$) as a function of number of users is shown in Fig. 3. The number of access points is also 20. It can be observed from Fig. 3 that our proposed user association strategy outperforms other three ones. Although the CRE scheme can obtain almost the same total rate requirement when bias is set to 12 dB, however, more users can access the H-CRAN by using our proposed user association strategy as illustrated in Fig. 2.

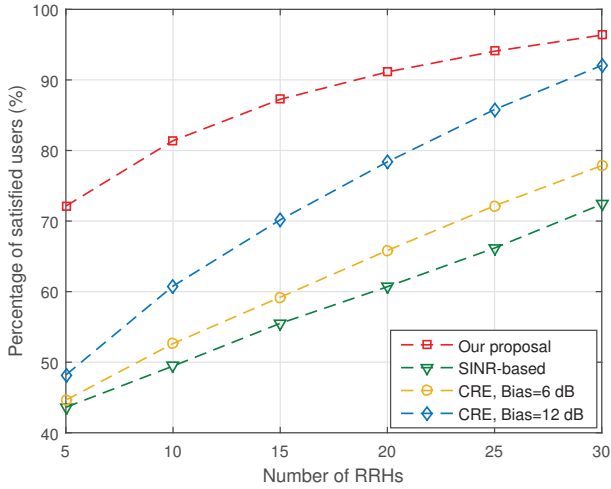


Fig. 4. Percentage of fully satisfied users as a function of the number of RRHs. $K = 200$.

Then we investigate the percentage of satisfied users with different number of RRHs. The number of users is 200. As shown in Fig. 4, our proposed user association strategy outperforms other three schemes in different scenarios. When the number of RRHs is small, i.e. radio resources of RRHs are strained, our proposal shows great improvement compared to others. The gap between our proposal and the CRE becomes smaller and smaller as the increasing of the number of RRHs. Based on the results in Fig. 2-4, we can conclude that the proposed user association strategy can utilize radio resources efficiently and effectively in different network scenarios.

V. CONCLUSIONS

In this paper, we studied the user association in the H-CRAN. First, we design a fast algorithm to obtain the minimum required power to satisfy a given number of users. Then, we introduce a *reference power* concept based on which an efficient user association algorithm is developed, which yields $\frac{1}{2}$ fraction of the optimum. Numerical results verified the effectiveness and efficiency of our proposal. For future work, dynamic RRHs switching on/off based on our proposed user association strategy is an interesting problem.

APPENDIX A PROOF OF LEMMA 1

Denote $p_{k,n}^*$ and $b_{k,n}^*$ as the optimal power and bandwidth allocation for $p_n(\mathcal{M}_1)$, respectively. Based on (16) and $p_n(\{k_1\}) \geq p_n(\{k_2\})$, $\forall k_1 \in \mathcal{M}_1, k_2 \in \mathcal{M}_2$, we have

$$\frac{1}{H_{k_1,n}} \cdot \left(2^{R_{k_1}^{min}/b_n^{max}} - 1 \right) \geq \frac{1}{H_{k_2,n}} \cdot \left(2^{R_{k_2}^{min}/b_n^{max}} - 1 \right). \quad (25)$$

Note the following mathematical fact: If $a_1 - 1 \geq A \cdot (a_2 - 1)$, where $a_1 > 1, a_2 > 1$, then $a_1^n - 1 \geq A \cdot (a_2^n - 1), \forall n \geq$

1, $A > 0$ holds. For each $k_1 \in \mathcal{M}_1, k_2 \in \mathcal{M}_2$, we have

$$\begin{aligned} p_{k_1,n}^* &= \frac{b_{k_1,n}^*}{H_{k_1,n}} \cdot \left(2^{\frac{R_{k_1}^{min}}{b_n^{max}} \cdot \frac{b_n^{max}}{b_{k_1,n}^*}} - 1 \right) \\ &\geq \frac{b_{k_1,n}^*}{H_{k_2,n}} \cdot \left(2^{\frac{R_{k_2}^{min}}{b_n^{max}} \cdot \frac{b_n^{max}}{b_{k_1,n}^*}} - 1 \right) \\ &= \frac{b_{k_1,n}^*}{H_{k_2,n}} \cdot \left(2^{R_{k_2}^{min}/b_{k_1,n}^*} - 1 \right), \end{aligned} \quad (26)$$

where the inequality follows by the fact that $b_{k_1,n}^*$ is always less than b_n^{max} . Equation (26) indicates that the power cost to serve users in \mathcal{M}_2 is less than $p_n(\mathcal{M}_1)$ even adopting the same bandwidth allocation.

Consider the same bandwidth allocation for the users in \mathcal{M}_2 , that is, each user $k_2 \in \mathcal{M}_2$ can get $b_{k_1,n}^*$ bandwidth, we have

$$\begin{aligned} p_n(\mathcal{M}_1) &= \sum_{k_1 \in \mathcal{M}_1} p_{k_1,n}^* \\ &\geq \sum_{k_2 \in \mathcal{M}_2} \frac{b_{k_1,n}^*}{H_{k_2,n}} \cdot \left(2^{R_{k_2}^{min}/b_{k_1,n}^*} - 1 \right) \\ &\geq p_n(\mathcal{M}_2). \end{aligned} \quad (27)$$

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [2] R. Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 94–101, May 2014.
- [3] S. Wang and C. Ran, "Rethinking cellular network planning and optimization," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 118–125, Apr. 2016.
- [4] China Mobile Research Institute, "C-RAN the road towards green RAN," Oct. 2011.
- [5] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126–135, Dec. 2014.
- [6] C. Ran, S. Wang, and C. Wang, "Balancing backhaul load in heterogeneous cloud radio access networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 42–48, June 2015.
- [7] J. Gong, J. S. Thompson, S. Zhou, and Z. Niu, "Base station sleeping and resource allocation in renewable energy powered cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3801–3813, Nov. 2014.
- [8] Y. Yang, L. Chen, W. Dong, and W. Wang, "Active base station set optimization for minimal energy consumption in green cellular networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5340–5349, Nov. 2015.
- [9] H. Tabassum, U. Siddique, E. Hossain, and M. J. Hossain, "Downlink performance of cellular systems with base station sleeping, user association, and scheduling," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5752–5767, Oct. 2014.
- [10] R. Madan, J. Borran, A. Sampath, N. Bhusan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1479–1489, Dec. 2010.
- [11] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," *IEEE Commun. Lett.*, vol. 15, no. 10, pp. 1084–1087, Oct. 2011.
- [12] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press: New York, 2004.
- [13] 3GPP, "Evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects (TR 36.814)," Mar. 2010.
- [14] —, "Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN (TR 36.932)," Mar. 2013.