Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Joint user association and resource partition for downlink-uplink decoupling in multi-tier HetNets*

Zhi-yong FENG[†‡1], Ze-bing FENG[1], T. Aaron GULLIVER[2]

([1]*MOE Key Laboratory of Universal Wireless Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China*)

([2]*Department of Electrical and Computer Engineering, University of Victoria, Victoria BC V8P 5C2, Canada*)

[†]E-mail: fengzy@bupt.edu.cn

Received Jan. 10, 2017; Revision accepted Mar. 25, 2017; Crosschecked June 4, 2017

Abstract: Traditional cellular networks require the downlink (DL) and uplink (UL) of mobile users (MUs) to be associated with a single base station (BS). However, the power gap between BSs and MUs in different transmission environments results in the BS with the strongest downlink differing from the BS with the strongest uplink. In addition, the significant increase in the number of wireless machine type communication (MTC) devices accessing cellular networks has created a DL/UL traffic imbalance with higher traffic volume on the uplink. In this paper, a joint user association and resource partition framework for downlink-uplink decoupling (DUDe) is developed for a tiered heterogeneous cellular network (HCN). Different from the traditional association rules such as maximal received power and range extension, a coalition game based scheme is proposed for the optimal user association with DUDe. The stability and convergence of this scheme are proven and shown to converge to a Nash equilibrium at a geometric rate. Moreover, the DL and UL optimal bandwidth partition for BSs is derived based on user association considering fairness. Extensive simulation results demonstrate the effectiveness of the proposed scheme, which enhances the sum rate compared with other user association strategies.

Key words: Downlink-uplink decoupling; User association; Resource partition; Heterogeneous cellular network; Coalition game
http://dx.doi.org/10.1631/FITEE.1700031 **CLC number:** TN929.5

1 Introduction

The performance of cellular networks has improved significantly in recent years. A heterogeneous deployment paradigm is now employed with a diversity of access points such as macrocells, microcells, picocells, femtocells, relay cells, and distributed antenna systems, with different capabilities and cell sizes coexisting in a geographic area (Ge *et al.*, 2016; Yang *et al.*, 2016). Dense deployment and spatial spectrum reuse have led to a tremendous increase in the sum rate of multi-tier heterogeneous cellular networks (HCNs). Due to the broadcast nature of wireless transmissions, mobile users (MUs) can receive signals from several overlaid base stations (BSs). Thus, flexible association with BSs can be employed to reduce downlink (DL) and uplink (UL) interference and improve throughput (Andrews, 2013).

1.1 Motivation

Multi-tier HCNs with randomly located small cell BSs provide MUs with numerous possible associations; i.e., an MU located on a macrocell cell edge

‡ Corresponding author

* Project supported by the National High-Tech R&D Program (863) of China (No. 2015AA01A705) and the National Natural Science Foundation of China (No. 61525101)

ORCID: Zhi-yong FENG, http://orcid.org/0000-0001-5322-222X

©Zhejiang University and Springer-Verlag Berlin Heidelberg 2017

may prefer to connect to a nearby femtocell BS to enhance signal quality and reduce intercell interference (López-Pérez *et al.*, 2011). Prior research in this area has considered the user association problem in HCNs from the perspective of outage probability (ElSawy *et al.*, 2013), maximizing the network throughput (Fooladivanda and Rosenberg, 2013), minimizing intercell interference (Madan *et al.*, 2010), minimizing BS power consumption (Son *et al.*, 2011), and cross-tier load balancing to enhance the network throughput (Singh *et al.*, 2014). However, these results are constrained by the traditional requirement that the downlink and uplink of a user are associated with the same BS. Due to the power gap between the BSs and MUs and different UL and DL transmission environments, there exists an imbalance in the downlink and uplink transmission conditions. As shown in Fig. 1, mobile users may have access to a strong downlink with the macro BS but the uplink to micro BS due to a close distance.

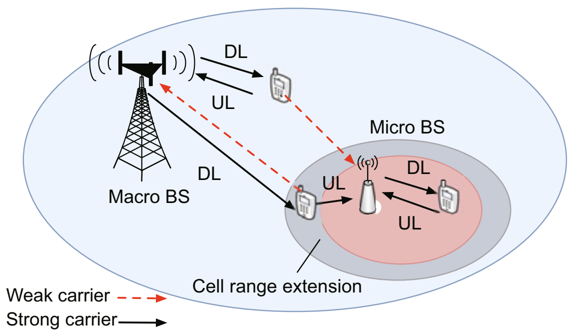

Fig. 1 A two-tier HCN composed of a macrocell and a microcell. Mobile users can establish downlink (DL) and uplink (UL) connections with BSs in different tiers

Therefore, a coupled downlink and uplink association may not provide both optimal DL and UL, which results in inefficient utilization of BS resources. Further, to ensure that a serving BS can recover the uplink signal, cell edge MUs with poor channel conditions must transmit at a high power level, which reduces battery life. As a consequence, it is better for MUs to retain the DL serving BS and establish an UL connection with another BS to improve performance.

Apart from the power gap, there is a significant imbalance in the DL and UL traffic load at a BS. MUs can have very different traffic characteristics, so the DL to UL traffic ratio can vary significantly. With the emergence of machine type communications (MTCs), uplink data rates can be much greater than downlink rates (Lien *et al.*, 2011). Millions of devices now access cellular BSs, and heavy loads on individual links (i.e., uplinks) result in resource allocation and scheduling issues at the BS and degradation in network performance. A solution to this problem is to allow MUs to establish an uplink connection with a BS that can provide a suitable signal quality and has a lighter load.

The imbalances described above can be addressed through flexible downlink and uplink associations that can guarantee high-quality links and more balanced traffic loads at the BSs. Fig. 2 shows the divergence of the user association in the DL and UL according to the maximum received signal to interference and noise ratio (SINR). This indicates that there is a significant imbalance in selecting a serving BS for both the DL and UL. Radio resource management can be conducted separately to satisfy both human and machine specific communication requirements. In this study, we focus on the downlink-uplink decoupling (DUDe) user association and the resource partitioning, and investigate the corresponding network performance improvement.

1.2 Related work

User association in multi-tier HCNs is essential to provide MUs with serving BSs. This is determined according to a set of rules and criteria. In Dhillon *et al.* (2012) and Jo *et al.* (2012), MUs are connected to BSs based on the maximum instantaneous SINR and the long-term average received power. The goal is to provide a high average transmission rate for all MUs. The maximum received power association rule is commonly employed. Cell range expansion and biasing were adopted in Ye *et al.* (2013) as an effective approach to expand cell coverage in order to offload MUs onto small cells. It was shown that this approach provides a throughput improvement compared with traditional association rules. Since BS resources are limited and cell loads differ between tiers, Andrews (2013) proposed a user-perceived rate that depends on both SINR and the load of the serving BS. User association and load balancing were jointly considered to formulate proportional fairness criteria for allocating MUs to BSs in different tiers. The goal with these approaches is to improve network performance by connecting MUs to serving BSs. However, MUs are associated with the same BSs for

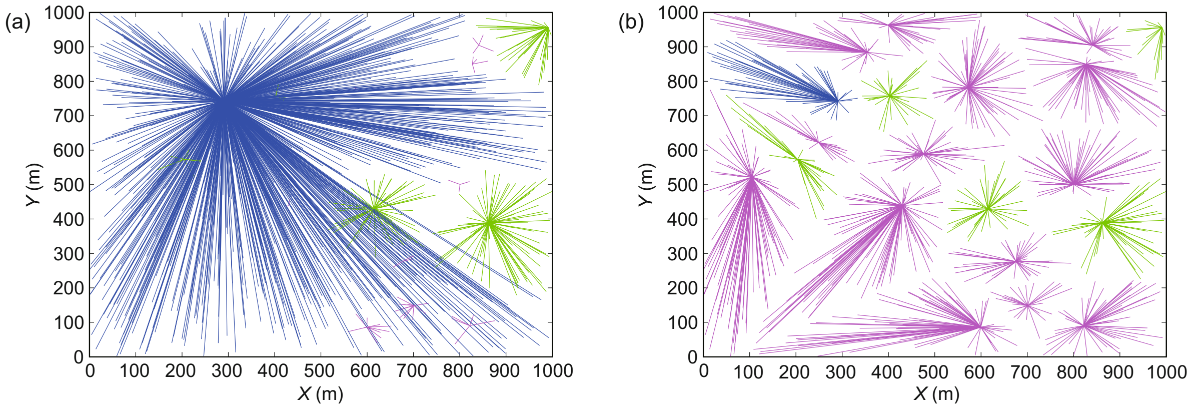

Fig. 2 User association in a three-tier heterogeneous network which includes microcell BSs (blue lines), picocell BSs (green lines), and femtocell BSs (purple lines) using the maximum received power level criteria: (a) downlink association; (b) uplink association. The MUs are located at the ends of the lines. References to color refer to the online version of this figure

both the DL and UL, which constrains the network performance.

Recently, DUDe has been considered as a promising framework for fifth generation (5G) cellular services. Boccardi *et al.* (2016) presented a detailed discussion on why the downlink and uplink should be decoupled from the perspectives of reducing transmission power, eliminating interference, improving uplink rate, balancing traffic load, and lowering deployment costs. They considered DUDe as an innovative approach to improving cellular network performance. In Elshaer *et al.* (2014), decoupled LTE downlink and uplink user association was considered based on the DL received power and UL path loss. Experiments were performed using the Vodafone two-tier LTE field trial network in London, and the results obtained show significant potential for uplink throughput enhancement. In Feng *et al.* (2014) and Smiljkovikj *et al.* (2015), decoupled DL/UL in a two-tier HCN was studied with macrocell BSs overlaid with femtocell BSs using a stochastic geometry framework. The results indicate that as the relative density of the femtocell BSs increases, more MUs associated with macrocell BSs prefer to decouple their uplink access to a femtocell BS. A joint rate and SINR coverage analysis was presented in Singh *et al.* (2015), which shows that a minimum path loss association is optimal for the uplink rate. These results indicate that employing different association rules for the downlink and uplink can significantly improve the system performance. The DUDe analysis was based on a stochastic Poisson point process (PPP) by averaging user SINR and rate. Jiang *et al.* (2016) studied the available resource allocation in two-tier HetNets with DUDe and formulated a non-linear integer programming model to solve the corresponding problem. However, the DUDe user association is significantly affected by the MU channel quality and available BS resources. In addition, DUDe can cause user load imbalance in the downlink and uplink, which requires an adjustment of the partition of spectrum resources to the downlink and uplink. Such issues are critical to the application of DUDe in HetNets but have not yet been investigated.

1.3 Contributions

In this paper, we propose a joint user association and resource partition framework for DUDe. The contributions of this work are given below:

1. We formulate the joint user association and resource partition optimization problem for DUDe as a 0-1 nonlinear integer programming problem. Using a suitable utility function, we transform the objective function into a convex form, which decouples the resource partition and user SINR parameters. Then the joint problem is separated into two subproblems, namely the user association problem (UAP) and resource partition problem (RPP).

2. A coalition game is proposed to model the UAP, and a user association coalition formation (UACF) algorithm is employed to solve the optimal user association problem. The stability and convergence of the algorithm are theoretically

proven, and the Nash equilibrium of coalition formation is approached. Then the optimal resource partition for downlink and uplink is derived considering utility fairness.

3. We evaluate the performance of the proposed scheme and compare it with other user association rules, which demonstrate the effectiveness and system utility improvement. Results are presented which show that the proposed scheme improves the system rate by 60% in DL and 80% in UL.

2 System model

We consider a general heterogeneous network model consisting of K tiers, e.g., macrocells, microcells, picocells, and femtocells. Generally, macrocell BSs have high transmission power to provide basic services having low to moderate rate requirements with a sufficient coverage for distributed mobile users. Low-power and low-cost small cells are deployed randomly as the second and/or third tier (typically overlaid with macrocells), for the purposes of offloading mobile users and offering additional capacity for uneven traffic requirements. Each tier has different parameters such as the BS transmission power, spatial density, and path loss. A mobile user can choose to access any BS in the network according to its transmission requirements and the network environment, e.g., minimum data rate, interference constraints, and available radio resources. The sets of BSs and MUs are denoted by B and C with cardinality $|B|$ and $|C|$, respectively. Throughout the paper, the superscripts D and U denote DL and UL parameters, respectively, and parameters without a superscript denote both DL and UL.

2.1 Propagation model

Free access is assumed so that an MU can associate with a BS in any tier according to some association rules, i.e., maximal received power level (max-RPL) (Jo *et al.*, 2012), range extension (Ye *et al.*, 2013), and minimum path loss (min-PL) (Singh *et al.*, 2015). A serving BS should always provide MUs with the maximum signal strength according to the association metrics, as otherwise they would shift to another BS. The received power at a node is

$$P_r = P_t |h_{t,r}|^2 d_{t,r}^{-\alpha}, \tag{1}$$

where P_t is the transmission power of transmitter t, P_r is the received power at receiver r, $d_{t,r}$ is the distance between transmitter t and receiver r, and α is the path loss exponent. We denote the channel fast fading random variable by $|h|^2$ with $E[|h|^2]=1$. The DL and UL channel SINR for MU i associated with serving BS j in tier k at a distance $d_{i,j}$ can then be expressed as

$$\gamma_{i,j}^{\mathrm{D}} = \frac{P_j |h_{i,j}^{\mathrm{D}}|^2 d_{i,j}^{-\alpha_{i,j}^{\mathrm{D}}}}{\sum_{j'\in B\backslash j} P_{j'} |h_{i,j'}^{\mathrm{D}}|^2 |d_{i,j'}|^{-\alpha_{i,j'}^{\mathrm{D}}} + \sigma^2}, \tag{2}$$

$$\gamma_{i,j}^{\mathrm{U}} = \frac{P_i |h_{i,j}^{\mathrm{U}}|^2 d_{i,j}^{-\alpha_{i,j}^{\mathrm{U}}}}{\sum_{i'\in C_I\backslash i} P_0 |h_{i',j}^{\mathrm{U}}|^2 |d_{i',j}|^{-\alpha_{i',j}^{\mathrm{U}}} + \sigma^2}, \tag{3}$$

where $j \in B\backslash j$ denotes the BSs not serving the MU in the downlink, and $i \in C_I\backslash i$ denotes the MUs not associated with BS j using the same resource block (RB) as user i in the uplink. P_j and P_0 are the transmission powers of BS j and the MU, respectively, and σ^2 is the additive white Gaussian noise (AWGN) power.

2.2 Downlink-uplink decoupling association framework

In a multi-tier heterogeneous cellular network (HCN), MUs can have multiple association options, for example, a high power macrocell BS supporting moderate data rates or a low power picocell BS located nearby which provides high capacity over short distances. This concept is extended here so that MUs can make separate decisions on which BSs to establish downlink and uplink associations. These decisions are based on parameters such as the channel conditions, traffic loads, and BS resources. DUDe implementation requires only that downlink and uplink control information be transmitted between the BSs and MUs. This can easily be achieved using network virtualization and network resource sharing (e.g., 3GPP RAN sharing). Fig. 3 shows the DUDe association and interference in a multi-tier HCN. The target MU has decoupled DL and UL with macrocell and picocell BSs, and suffers from co-channel DL and UL interference from other tier BSs.

We consider that in general DUDe association follows the preferable BSs for user access in different tiers. We assume that BSs in the kth tier are distributed as a homogeneous Poisson point process (PPP) with density λ_k and transmission power

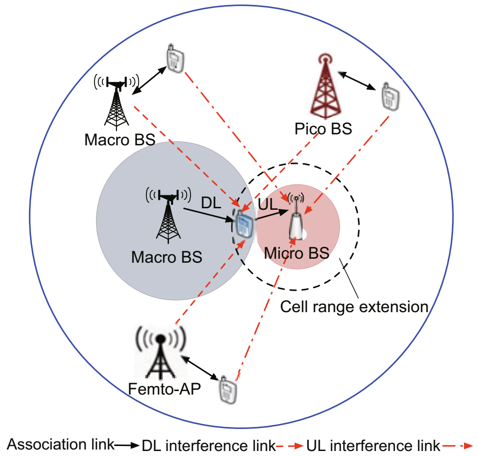

Fig. 3 Downlink-uplink decoupling association and interference in multi-tier heterogeneous cellular networks

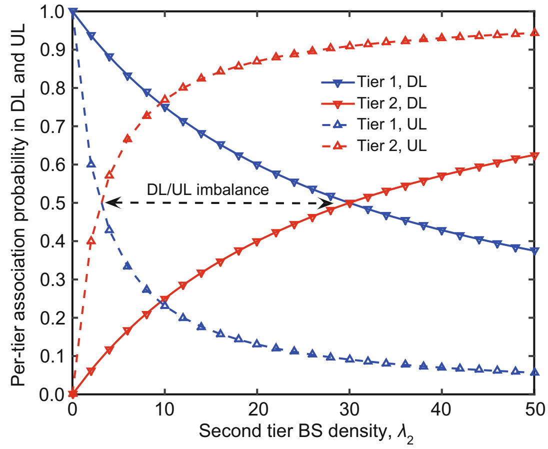

Fig. 4 Per-tier user association probability with $\lambda_1 = 3$

P_k. Moreover, MUs are considered to be distributed across the region with a PPP distribution independent of the BSs with density λ_0 and the same transmission power. Based on our previous work (Feng *et al.*, 2014), the downlink and uplink association probabilities for a specific MU with different tiers following the max-RPL rule are

$$\nabla_k = 2\pi\lambda_k \int_0^\infty r \exp\left(-\sum_{j=1}^K \pi\lambda_j \left(\frac{P_j}{P_k}\right)^{\frac{2}{\alpha_j}} r^{\frac{2\alpha_j}{\alpha_k}}\right) \mathrm{d}r, \tag{4}$$

$$\Delta_k = 2\pi\lambda_k \int_0^\infty r \exp\left(-\pi \sum_{j=1}^K \lambda_j r^{\frac{2\alpha_j}{\alpha_k}}\right) \mathrm{d}r. \tag{5}$$

Assuming that all tiers have the same path loss, the results can be simplified to obtain the user association imbalance in tier k as

$$\rho_k = \frac{\Delta_k}{\nabla_k} = \frac{\sum\limits_{j=1}^K \lambda_j (P_j/P_k)^{2/\alpha_j}}{\sum\limits_{j=1}^K \lambda_j}. \tag{6}$$

This indicates that the BS downlink and uplink load imbalance is proportional to the transmission power ratio P_j/P_k. Fig. 4 presents the per-tier association probability in a two-tier HetNet (first tier macrocells and second tier microcells). We observe that the MUs have an asymmetric association preference in the DL and UL in both tiers. Further, as the second tier BS density increases, more MUs are offloaded from macrocells to microcells. This motivates investigating the DUDe optimal user association problem, which is formulated in the next subsection.

2.3 Problem formulation

With traditional user association, the connection between an MU and a BS can be denoted as a binary value 0 or 1 representing whether the MU is associated with the BS or not, and the optimization is confined to the downlink. With DUDe, the downlink and uplink user association optimization is divided. Let $x_{ij}^{\mathrm{D}} = 1$ denote that user i is associated with BS j in the DL and $x_{ij}^{\mathrm{D}} = 0$ otherwise. The corresponding variable in the UL is x_{ij}^{U}. We assume that an MU is associated with one BS in the downlink and one BS in the uplink, so there are three cases, i.e., downlink-only associated, uplink-only associated, or both downlink and uplink associated. Note that all MUs must be associated with service BSs, i.e., $\sum_{j\in B} x_{ij}^{\mathrm{D}} = \sum_{j\in B} x_{ij}^{\mathrm{U}} = 1$. The network association map matrices in the downlink and uplink are denoted as $\boldsymbol{X}^{\mathrm{D}} = \left\{x_{ij}^{\mathrm{D}} \in \{0,1\} \middle| i \in C, j \in B\right\}$ and $\boldsymbol{X}^{\mathrm{U}} = \left\{x_{ij}^{\mathrm{U}} \in \{0,1\} \middle| i \in C, j \in B\right\}$, respectively.

Typically, proportional fairness is chosen as the objective function (Fooladivanda and Rosenberg, 2013), i.e., maximizing the utility function $U = \sum_{i\in C} \ln \lambda_i$ where λ_i is the throughput of user i, which is closely related to the user SINR and the allocated RBs. A function that maps the user downlink/uplink SINR and allocated RBs to the corresponding data rate can then be employed. For simplicity, we consider channel capacity as the mapping function. The downlink instantaneous rate of user i served by BS j is then

$$r_{ij} = \frac{W\eta}{\sum_{i\in C} x_{ij}} \log(1+\gamma_{ij}), \tag{7}$$

where W is the total bandwidth and equal partitioning of the downlink (uplink) resources among the associated users is assumed (Singh *et al.*, 2015). This is not only a suitable rate expression but is also a convex function which can easily be evaluated. With DUDe, the above expressions can be separated, so the DL and UL throughput of user i can be expressed as $\lambda_i^{\mathrm{D}} = \sum_{j\in B} x_{ij}^{\mathrm{D}} r_{ij}^{\mathrm{D}}$ and $\lambda_i^{\mathrm{U}} = \sum_{j\in B} x_{ij}^{\mathrm{U}} r_{ij}^{\mathrm{U}}$.

The optimization problem is to determine $\boldsymbol{X}^{\mathrm{D}}$ ($\boldsymbol{X}^{\mathrm{U}}$) and the DL/UL resource partitioning fraction η to maximize the overall downlink (uplink) utility. Thus, we can formulate the joint user association and resource partition problem for DUDe as

$$P_{\mathrm{DL}}: \max_{\boldsymbol{X}^{\mathrm{D}},\eta^{\mathrm{D}}} \sum_{i\in C} \ln(\lambda_i^{\mathrm{D}}) \tag{8a}$$

$$\text{s.t.} \ \sum_{j\in B} x_{ij}^{\mathrm{D}} = 1, \quad i \in C, \tag{8b}$$

$$x_{ij}^{\mathrm{D}} \in \{0,1\}, \quad i\in C,\ j\in B, \tag{8c}$$

and

$$P_{\mathrm{UL}}: \max_{\boldsymbol{X}^{\mathrm{U}},\eta^{\mathrm{U}}} \sum_{i\in C} \ln(\lambda_i^{\mathrm{U}}) \tag{9a}$$

$$\text{s.t.} \ \sum_{j\in B} x_{ij}^{\mathrm{U}} = 1, \quad i \in C, \tag{9b}$$

$$x_{ij}^{\mathrm{U}} \in \{0,1\}, \quad i\in C,\ j\in B. \tag{9c}$$

This is an integer programming problem and is thus NP-hard. Moreover, the utility function is not increasing or concave with respect to x_{ij}^{D} or x_{ij}^{U}. However, an upper bound can be obtained by relaxing the integer constraints on $x_{ij}^{\mathrm{D}}, x_{ij}^{\mathrm{U}}$ as $0 \le x_{ij}^{\mathrm{D}}, x_{ij}^{\mathrm{U}} \le 1$. Thus, the original objective function can be reformulated to a convex form:

$$\begin{aligned}
\sum_{i\in C} \ln\lambda_i &= \sum_{j\in C} \ln\left(\sum_{j\in B} x_{ij} r_{ij}\right) \\
&\overset{(a)}{=} \sum_{i\in C}\sum_{j\in B} x_{ij} \ln r_{ij} \\
&= -\sum_{i\in C}\sum_{j\in B} \ln\left(\sum_{i\in C} x_{ij}\right) \\
&\quad + \sum_{i\in C}\sum_{j\in B} x_{ij}\left\{ \ln(W\eta_j) + \ln[\log(1+\gamma_{ij})] \right\},
\end{aligned} \tag{10}$$

where (a) follows from the binary characteristics of x_{ij}^{D}. This shows that the utility is a function of both resource partition fraction η and user association matrix $\boldsymbol{X}$. Given that the expression is the sum of several terms, we can split the joint optimization problem into two subproblems, namely the user association problem (UAP) and the resource partition problem (RPP). Since the user SINR is unrelated to η, the optimal downlink and uplink user association can be derived first independent of η. Based on the association results, the RPP solution is then obtained considering downlink and uplink utility fairness. Since the resulting objective function is convex, the UAP becomes a convex integer optimization problem. Therefore, the upper bound of the relaxed problem can be solved even for large networks in polynomial time (Nesterov and Nemirovski, 1994). However, the association mapping matrix $\boldsymbol{X}$ can be very large given the number of MUs and BSs in the network, and the solution space grows exponentially with the number of BSs and MUs. In the next section, a coalition game based algorithm is proposed to solve the user association problem. Then based on the association results, the optimal resource partition in the downlink and uplink is derived considering utility fairness.

3 Downlink-uplink decoupling user association optimization

3.1 Coalition formation game

In the coalition game, MUs act as players with the goal of improving system utility. Since MUs choose BSs to associate with and share the BS resources for transmission, we can develop $|B|$ coalitions $\{G_1, G_2, ..., G_{|B|}\}$, and in G_b the MUs are players associated with BS b, denoted as $G_b = \{i | x_{ib} = 1, i \in C, b \in B\}$. In this game, players tend to form or join coalitions to maximize their individual utilities. Considering the max-RPL association rule, if more MUs merge into one coalition with the maximum RPL, then the average number of allocated RBs for each MU decreases due to the limited spectrum resources, which reduces the overall coalition utility. Thus, merging into one coalition is not advantageous, so MUs will possibly deviate from each other and join different coalitions. Nevertheless, the coalition form is closely related to the coalition utility, which is affected by user association. Before proposing the coalition formation solution, we follow the approach in Saad *et al.* (2011) and define the user association coalition game structure.

Definition 1 The user association coalition game for MUs associated with BSs in the downlink and uplink can be defined by a triple (C, G, U), and the game formulation is as follows:

1. Players: MUs are game players with user set C.

2. Transferable utility: We denote $U(G_b) \triangleq \sum_{i \in G_b} \ln \lambda_i$ as the transferable utility of coalition G_b, which denotes the total value of the coalition. Considering the payoff for each user in the coalition, $U(G_b)$ is composed of the individual contribution of all members. We define the individual utilities by distributing the sum utility among all members of the coalition as

$$U_i(G_b) = \frac{\ln \lambda_i}{\sum_{j \in G_b} \ln \lambda_j}. \tag{11}$$

Note that each MU should have the channel state information of the other MUs in the same coalition to calculate its own utility. Since all MUs in the same coalition G_b are associated with the same BS b, centralized management can easily be used to measure the channel conditions and broadcast this information to the MUs.

3. Coalition partition: The user association coalition partition is defined as $G = \{G_1, G_2, \cdots, G_{|B|}\}$, where $G_i \cap G_{i'} = \varnothing$ for $i \neq i'$ and $\cup_{i=1}^{|B|} G_i = C$.

4. Strategy: The strategy for the players is to determine the association matrix $\boldsymbol{X}$ and select which coalition to form according to the utility of the coalition as well as other coalitions.

5. Nash equilibrium: We denote the coalition partition $G = \{G_1, G_2, \cdots, G_{|B|}\}$ as the Nash equilibrium if $\forall\, i \in C$ and $i \in G_b$, $G_b \succ_i G_{b'}$ for all $G_{b'} \in G$, where $\succ_i$ denotes the preference order.

3.2 Coalition formation algorithm

The key strategy for coalition formation is to enable players to associate with the coalition according to the coalition utilities. In a distributed manner, each MU should compare and order its own potential utility to decide which coalition to join. From Eq. (10), the total system utility is finite. Thus, the player coalitions can be changed in a finite number of iterations to improve the system utility. The order of player preferences is given in the following definition:

Definition 2 The preference order $\succ_i$ is defined as a complete, reflexive, and transitive binary relation over the set of all coalitions that MU i can form.

With this definition, $\succ_i$ can be used to describe the preference of MU i to join a coalition. We denote $G_b \succ_i G_{b'}$ as the preference of MU i to be a member of coalition G_b rather than $G_{b'}$; in other words, MU i prefers to associate with BS b rather than b'. Since the preference order depends on the MU utility, we define that a preference change occurs when

$$G_b \succ_i G_{b'} \Leftrightarrow \begin{cases} U_i(G_b) > U_i(G_{b'}),\ i \in G_b, \\ U_j(G_l) \geq U_j(G_l),\ j \in G_l,\ j \neq i. \end{cases} \tag{12}$$

From the definition, we know that MU i prefers to leave the current coalition $G_{b'}$ and join the new coalition G_b if and only if two conditions are satisfied. First, the coalition change increases the utility of MU i. Second, no MUs in other coalitions suffer from a utility decrease due to this coalition change. Therefore, the overall system utility is not optimum if there exists a preference order $G_b \succ_i G_b'$, and thus the association is not optimum. Based on the MU coalition preferences, we propose a search of the system utility by switching MU coalitions. Given a current association matrix $\boldsymbol{X}$ and coalition G, MU i is switched from G_b to $G_{b'}$. After the switch, the coalition formation is $\{G_1, G_2, \cdots, G_b \backslash \{i\}, \cdots, G_{b'} \cup \{i\}, \cdots, G_{|B|}\}$. Via repeated switch operations, different coalitions are formed denoting the MU associations with preferred BSs. According to the preference order defined in Eq. (12), the switch for MU i is kept if and only if $G_b \succ_i G_b'$. Then the system utility is improved via MU coalition switch operations, and the maximum system utility can be obtained using the proposed user association coalition formation algorithm presented in Algorithm 1.

Before performing switch operations, we use conventional association rules to initialize the coalition partitions, i.e., max-RPL and min-PL rules. Due to the random locations of MUs and BSs, the number of MUs connected with a BS may be very large under the user association rules, which causes cell overload and affects user utility. In the core iterative procedure, the algorithm randomly chooses an MU and switches its association to another possible coalition. To determine whether the switch is kept, the coalition utility after the switch is calculated. For downlink user association, an MU coalition switch from G_b to $G_{b'}$ does not affect other coalition utilities because the MU downlink interference from the BSs does not change. Thus, only the utility of two

Algorithm 1 User association coalition formation algorithm in the uplink and downlink

```
1: Initialize user association matrix X^ini according to the predefined user association rules;
2: Form the current coalition partition as G^cur based on X^ini;
3: repeat
4:    Uniformly randomly select one MU i ∈ C and determine coalition G_b;
5:    Uniformly randomly select another coalition G_b' ∈ G^cur, G_b' ≠ G_b;
6:    Calculate U(G_i), i = 1, 2, ···, |B|;
7:    SwitchFlag = 0;
8:    if G_b' ≻_i G_b then
9:       SwitchFlag = 1;
10:   else
11:      Calculate the probability
         θ_{b,b'}(T_n) = exp((U(G_b') − U(G_b))/T_n);
12:      Randomly generate β uniformly distributed in (0, 1];
13:      if β < θ_{b,b'}(T_n) then
14:         SwitchFlag = 1;
15:      end if
16:   end if
17:   if SwitchFlag = 1 then
18:      MU i leaves the current coalition G_b and joins the new coalition G_b';
19:      Update the current coalition partition set as
         (G^cur\{G_b, G_b'}) ∪ (G_b\{i}) ∪ (G_b' ∪ {i}) → G^cur;
20:   end if
21: until user association converges to the final Nash-stable coalition partition G^final
22: return user association matrix X^final with elements x_ij = 1, i ∈ G_j, G_j ∈ G^final;
```

coalitions G_b and $G_{b'}$ is required to decide whether or not to keep the switch according to Eq. (12). The MU association switch will affect the occupied resources in the UL, so the interference to other MUs will change. However, user offloading usually occurs among different overlapping tiers and the inter-tier interference is more dominant than the intra-tier interference (Dhillon *et al.*, 2013). Hence, the MU utility for the two coalitions, instead of for all MUs, can be used in the system, which results in a much lower complexity compared with the centralized solution.

The randomly chosen switch operations may result in a locally optimal solution. To avoid this situation, with a given probability the switch is kept when the new coalition does not follow the preference order. This probability is used in step 11 of Algorithm 1 and is defined as

$$\theta_{b,b'}(T_n) = \exp\left(\frac{U(G_{b'}) - U(G_b)}{T_n}\right), \tag{13}$$

where $T_n = T_0/\log(n-1)$ with T_0 a constant and n the number of switch operations. Then the switch occurs with probability $\beta < \theta_{b,b'}(T_n)$ as the Barker sampler, which can enlarge the feasible space of the user association solutions (Brémaud, 1999). This condition can prevent the solution from deviating from the globally optimal solution, and encourages each MU to form a coalition that improves the system utility. After a finite number of switch iterations, the final stable coalitions denote the globally optimal user association. The stability and convergence of the UACF algorithm are analyzed in the next subsection.

3.3 Stability and convergence

We investigate the stability property of the UACF algorithm using the concept of Nash equilibrium in Definition 1. If there exists a Nash equilibrium, a user coalition switch will not be performed and the algorithm is stable. Then the stability and convergence properties of the UACF algorithm are given in the following theorem:

Theorem 1 The proposed user association coalition formation algorithm will always converge to a Nash equilibrium coalition partition G^{final} with probability 1 through a finite number of iterations.

Proof Since the total number of coalition partitions is equal to the number of BSs $|B|$, MU i can make only $|B|-1$ possible switches. As the number of iterations increases, the system utility also increases based on the coalitions formed. Moreover, from Eq. (13) we have

$$\begin{aligned}\lim_{n\to\infty}\theta_{b,b'}(T_n) &= \lim_{T_n\to 0}\exp\left(\frac{U(G_{b'}) - U(G_b)}{T_n}\right)\\ &= 0, \qquad \text{if } U(G_{b'}) < U(G_b).\end{aligned} \tag{14}$$

In this case, the switch operation that reduces the system utility is adopted with probability 0 because the condition in step 13 is no longer satisfied. Thus, the random switch operations will terminate with probability 1, and the system utility converges to a final coalition partition G^{final} after a finite number of iterations.

We prove stability via reduction to absurdity. If G^{final} is not stable, then there must exist an

MU $i \in G_b$ and a coalition $G_{b'} \neq G_b$ such that $G'_b \cup \{i\} \succ_i G_b$. According to the UACF algorithm, a switch operation for i can be performed to improve the coalition utility as $U(G_{b'} \cup \{i\}) > U(G_b \backslash \{i\})$. This contradicts the fact that G^{final} is the final coalition partition after convergence. Thus, the coalition partition G^{final} and the corresponding user association matrix $\boldsymbol{X}^{\text{final}}$ are in Nash equilibrium and stable, which completes the proof.

Theorem 1 proves that the proposed algorithm converges to a stable coalition partition after a finite number of iterations. The convergence rate is given in the following theorem:

Theorem 2 The convergence rate of the proposed UACF algorithm to the final Nash equilibrium user association is geometric.

Proof As the UACF algorithm applies Barker's sampler to accept the switch operation with probability $\theta_{b,b'}(T_n)$, we can use a Markov chain to describe the coalition formation state transition process. Denoting the initial user association state as $\boldsymbol{v}$ and the system transition matrix as $\boldsymbol{P}$, the distance between the system state after switch iteration n and the final optimal state $\boldsymbol{X}$ is $d(\boldsymbol{vP}^n, \boldsymbol{X})$. Following Definition 7.1 in Chapter 6 in Brémaud (1999), we then have

$$d(\boldsymbol{vP}^n, \boldsymbol{X}) \leq \delta(\boldsymbol{Q})\delta(\boldsymbol{P})^n = \frac{1}{2} d(\boldsymbol{v}, \boldsymbol{X})\,\delta(\boldsymbol{P})^n, \tag{15}$$

where $\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{X} \end{pmatrix}$, and $\delta(\boldsymbol{Q})$ is the ergodic coefficient of matrix $\boldsymbol{Q}$, defined as

$$\delta(\boldsymbol{Q}) = \frac{1}{2} \sup \sum_{k=1}^{|\boldsymbol{v}|} |q_{1k} - q_{2k}|. \tag{16}$$

Considering the characteristics of the ergodic coefficient and the transition probability in Eq. (13), we have

$$\delta(\boldsymbol{P}) \leq 1 - \exp\left(\frac{U(G_{b'}) - U(G_b)}{T_n}\right), \tag{17}$$

and therefore

$$d(\boldsymbol{vP}^n, \boldsymbol{X}) \leq \frac{1}{2} d(\boldsymbol{v}, \boldsymbol{X}) \left(1 - \mathrm{e}^{(U(G_{b'}) - U(G_b))/T_n}\right)^n. \tag{18}$$

Thus, the convergence from the initial state $\boldsymbol{X}^{\text{ini}}$ to the final optimal and stable state $\boldsymbol{X}^{\text{final}}$ is geometric, which completes the proof.

The performance of the UACF algorithm depends on the number of iterations. Thus, as the number of switches is increased, the algorithm will provide better user associations and a higher system utility. However, even though the number of iterations affects the final association matrix $\boldsymbol{X}^{\text{final}}$, according to Theorem 2 in Li *et al.* (2014), the algorithm converges to the optimal solution, and this is confirmed by the simulation results.

4 Downlink-uplink decoupling resource partition optimization

The UAP was solved above via a coalition game, which provides the optimal user association matrix $\boldsymbol{X}$. We now analyze the downlink and uplink resource partition problem based on this association. First, Eq. (10) can be expressed as

$$\sum_{i \in C} \ln \lambda_i = \sum_{j \in B} a_j \ln(W\eta_j) + a_2, \tag{19}$$

where a_j is the number of MUs associated with BS j and a_2 is a constant. We observe that Eq. (19) is a linear-log increasing function with resource partition fraction η. Since $\eta_j^{\mathrm{D}} + \eta_j^{\mathrm{U}} = 1$ for $j \in B$, the utility maximization for the downlink and uplink is a tradeoff; i.e., increasing the downlink (uplink) spectrum resources will reduce the uplink (downlink) system utility. A general method to solve this problem is to apply fairness so that the resource partition balances the downlink and uplink utility. For BS j, we have the following inequality:

$$\begin{aligned} & a_j^{\mathrm{D}} \ln(W\eta_j^{\mathrm{D}}) + a_j^{\mathrm{U}} \ln(W\eta_j^{\mathrm{U}}) \\ & = a_j^{\mathrm{D}} \ln(W\eta_j^{\mathrm{D}}) + a_j^{\mathrm{U}} \ln(W(1-\eta_j^{\mathrm{D}})) \\ & \overset{(a)}{\leq} \ln\left(W \frac{{a_j^{\mathrm{D}}}^{a_j^{\mathrm{D}}} {a_j^{\mathrm{U}}}^{a_j^{\mathrm{U}}}}{\left(a_j^{\mathrm{D}} + a_j^{\mathrm{U}}\right)^{\left(a_j^{\mathrm{D}} + a_j^{\mathrm{U}}\right)}}\right), \end{aligned} \tag{20}$$

where (a) uses the extreme value of the function when $\eta_j^{\mathrm{D}} = a_j^{\mathrm{D}}/(a_j^{\mathrm{D}} + a_j^{\mathrm{U}})$ $(\eta_j^{\mathrm{U}} = a_j^{\mathrm{U}}/(a_j^{\mathrm{D}} + a_j^{\mathrm{U}}))$. Using this resource partition rule, the sum utility of the downlink and uplink reaches its maximum.

The solution obtained is a long-term solution. Since an equal partition resource allocation has been applied (which can be realized by proportional fairness or round robin scheduling), MUs may occupy different RBs, which results in a complex interference map. The downlink interference is not affected by the resource allocation because the interference is from the BSs. However, for the uplink, the interference at a specific MU from other MUs is related to BS resource scheduling. Thus, changes in η may affect uplink user association due to changes in uplink interference. This issue does not exist in LTE-like systems because single carrier frequency division multiple access (SC-FDMA) is employed in the uplink. Moreover, inter-cell interference coordination (ICIC) can be used to schedule the uplink RBs to minimize the co-channel interference among MUs, which reduces the influence of resource partition on user association. Even though our solution is for long term, the proposed approach can be extended to obtain a suitable realization in practical networks. In this case, user association and DUDe resource partition can be employed alternately.

5 Performance results

Based on the model and analysis presented in the previous sections, the network performance with DUDe is now examined. The simulation environment is a $500\,\text{m} \times 500\,\text{m}$ area consisting of a two-tier HCN (macrocell tier and microcell tier) with random uniformly distributed BSs and MUs. Monte Carlo simulation is employed to evaluate the system performance. User association and resource partition results using the proposed algorithm are also given. The transmission power of the network tiers is $P_1 = 46$ dBm and $P_2 = 33$ dBm, and the thermal noise power is -174 dBm/Hz. The downlink and uplink path loss exponents are set to 4.

To verify the advantages of the UACF algorithm, we compare the performance of the proposed scheme with the following user association rules: (1) optimal solution obtained by an exhaustive search of all coalitions to find the association that provides the maximum system utility; (2) max-RPL rule (Jo *et al.*, 2012), where an MU is associated with the BS providing the highest received power level; (3) min-PL rule (Singh *et al.*, 2015), where an MU is associated with the BS with the minimum path loss; (4) biased-RPL rule (Ye *et al.*, 2013), where a bias factor B for different tiers is added to the max-RPL rule. Note that the above rules (2), (3), and (4) are the same for uplink association (no bias for uplink), so min-PL is used to evaluate uplink association.

5.1 Average system performance

Figs. 5a and 5b present the overall downlink and uplink network rate versus the number of microcell BSs deployed. These results show that the total rate increases as the number of microcell BSs increases. The reason is that more BSs can improve spectrum resource reuse. The min-PL rule performs better than the max-RPL rule because MUs are offloaded to the second tier cells. In this case, cell edge users with a low SINR are associated with microcells, which enhances macrocell user performance and efficiently uses microcell spectrum resources. The performance of the biased-RPL scheme is superior to that of max-RPL, but inferior to that of min-PL, with a higher system rate for a large bias factor B_2. In the downlink, the min-PL rule provides better performance than the non-DUDe technique (Singh *et al.*, 2015; Boccardi *et al.*, 2016). However, the proposed UACF algorithm provides a 60% improvement in the system rate over the other association schemes. This is because user association and load balance are jointly considered for each BS.

Figs. 6a and 6b show the overall downlink and uplink network rate versus the number of MUs. The overall rate increases because a greater number of MUs can more efficiently use the network resources and this improves the aggregate rate. When the number of MUs is 50, the UACF scheme provides a rate 80% better than with the other association rules. This improvement increases with the number of MUs in the network. This is because the proposed scheme can appropriately offload MUs to different BSs instead of just one BS, which avoids cell overload. The results in Figs. 5 and 6 show that the UACF algorithm provides the best performance, which confirms the optimality of the coalition formation procedure.

Figs. 7a and 7b present the downlink and uplink load in a two-tier HCN. As the number of microcell BSs increases, MU preference to associate with microcell BSs rather than macrocell BSs increases in both the downlink and uplink. This indicates the advantage of offloading to small cells in HetNets. Nevertheless, downlink and uplink imbalance still exists due to the separate DUDe association. In terms of the number of MUs in the network, we note that the load of the microcell tier is relatively constant in the uplink. This is because the proposed scheme takes user fairness into account and determines the optimal user association from the system rate perspective.

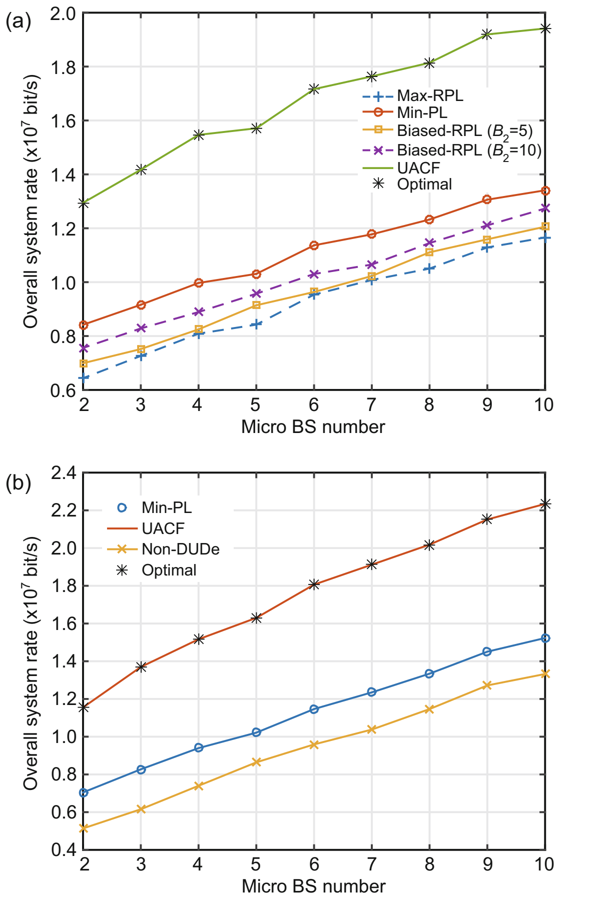

Fig. 5 System rate versus the number of microcell BSs in the downlink (a) and uplink (b) with different association schemes. There are three macrocell BSs and 30 MUs, and the bias factor for the macrocell tier is $B_1 = 1$

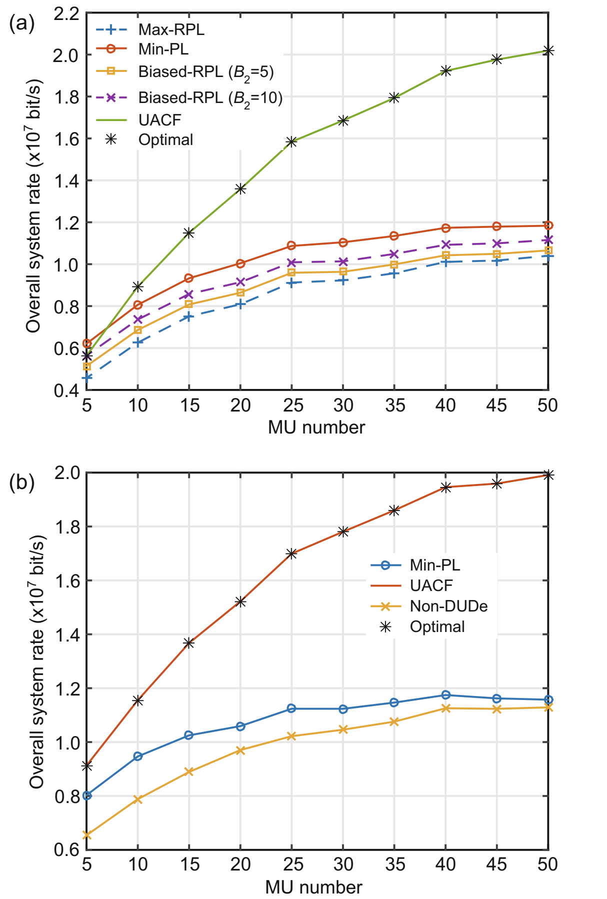

Fig. 6 System rate versus the number of MUs in the downlink (a) and uplink (b) with different association schemes. There are three macrocell BSs and six microcell BSs, and the bias factor for the macrocell tier is $B_1 = 1$

5.2 User association

To further examine the proposed scheme, we consider user association and resource partition. Figs. 8a and 8b present the user load in different tiers based on user association. These show that all association schemes offload users from macrocells to microcells. The min-PL rule has more offloading than the max-RPL and biased-RPL rules because users with a large macrocell BS RPL will associate with nearby microcell BSs. Further, the proposed scheme provides the best user offloading to microcell BSs as it has the highest network rate. A similar result is observed in the uplink as shown in Fig. 8b. Overall, the proposed DUDe user association scheme can effectively use the spectrum resources of the second tier small cells, and offload users to improve the system throughput.

The optimal resource partition for different macrocell and microcell BSs is presented in Fig. 9. This shows that the BSs have different fractions of resources allocated to the downlink and uplink. The macrocell BSs (1, 2, and 3) have more resources on the downlink than on the uplink, which means MUs prefer to download data from a macrocell BS that provides good channel quality. In the uplink, microcells are important as they can be used to offload MU data.

6 Conclusions

In this paper, joint user association and resource partitioning for downlink and uplink decoupling were investigated for a multi-tier heterogeneous cellular network (HCN). With decoupling,

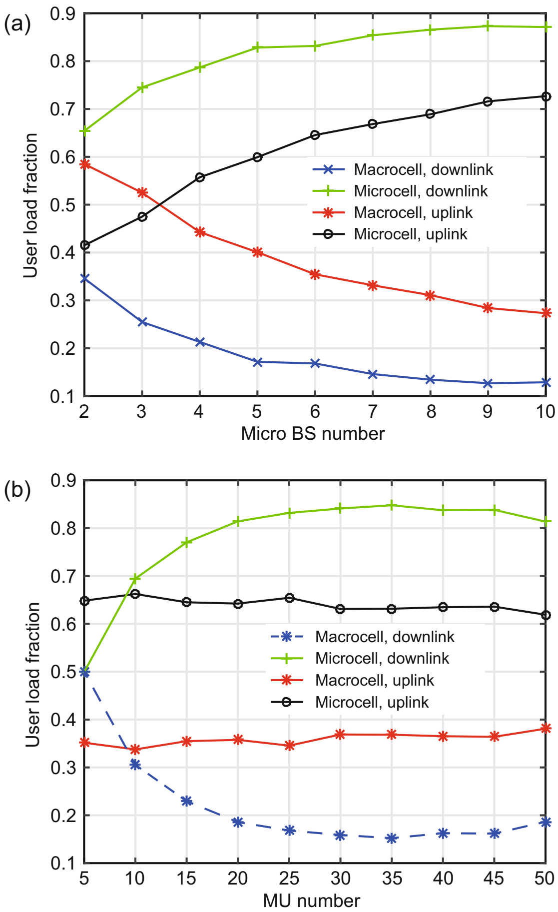

Fig. 7 Downlink and uplink load fractions with the proposed UACF scheme: (a) user load fraction versus the number of microcell BSs; (b) user load fraction versus the number of MUs

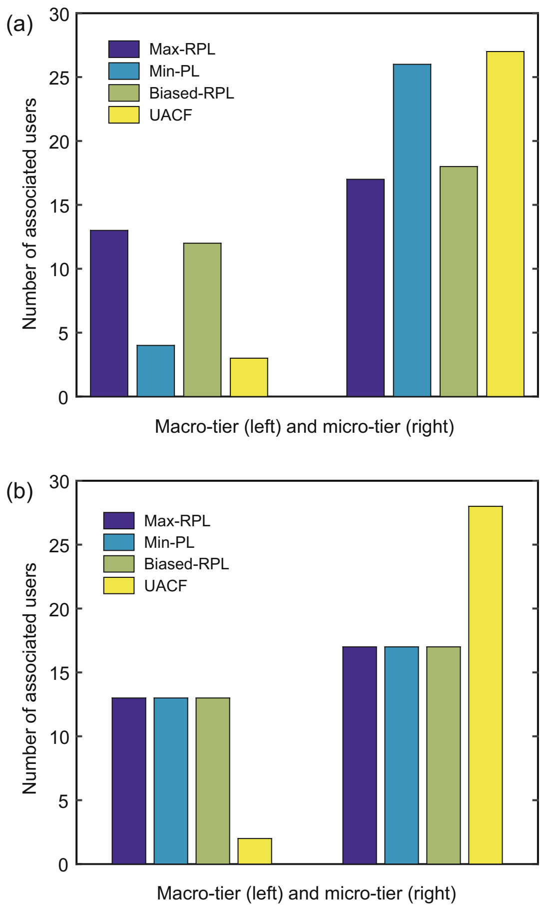

Fig. 8 Two-tier HCN user association in the downlink (a) and uplink (b)

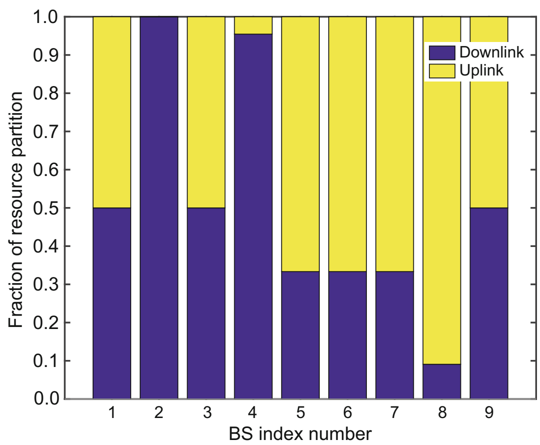

Fig. 9 The downlink and uplink resource partition fraction for each BS

mobile users (MUs) can separately associate with base stations (BSs) to maximize system utility, and the BS spectrum resources can be partitioned between the downlink and uplink based on the load. The joint problem was divided into a user association problem and a resource partition problem. A coalition game was proposed to obtain the optimal DL and UL association, and the stability and convergence of the proposed scheme were proven. Moreover, the optimal resource partition was derived considering MU utility fairness. Performance results were presented which demonstrated the effectiveness of the proposed approach; in particular, the network rate was improved by 60% compared with those of other schemes. Thus, DUDe is a promising framework to satisfy the requirements of fifth generation (5G) cellular networks. Future work will consider asymmetrical DUDe resource allocation such as flexible resource scheduling models for MUs to reduce inter-tier interference and improve system throughput.

References

Andrews, J.G., 2013. Seven ways that HetNets are a cellular paradigm shift. *IEEE Commun. Mag.*, **51**(3):136-144. http://dx.doi.org/10.1109/MCOM.2013.6476878

Boccardi, F., Andrews, J., Elshaer, H., *et al.*, 2016. Why to decouple the uplink and downlink in cellular networks and how to do it. *IEEE Commun. Mag.*, **54**(3):110-117. http://dx.doi.org/10.1109/MCOM.2016.7432156

Brémaud, P., 1999. Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues. Springer-Verlag, New York, NY.

Dhillon, H.S., Ganti, R.K., Baccelli, F., *et al.*, 2012. Modeling and analysis of k-tier downlink heterogeneous cellular networks. *IEEE J. Sel. Areas Commun.*, **30**(3):556-560. http://dx.doi.org/10.1109/JSAC.2012.120405

Dhillon, H.S., Ganti, R.K., Andrews, J.G., 2013. Load-aware modeling and analysis of heterogeneous cellular networks. *IEEE Trans. Wirel. Commun.*, **12**(4):1666-1677. http://dx.doi.org/10.1109/TWC.2013.13.120485

ElSawy, H., Hossain, E., Haenggi, M., 2013. Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: a survey. *IEEE Commun. Surv. Tutor.*, **15**(3):996-1019. http://dx.doi.org/10.1109/SURV.2013.052213.00000

Elshaer, H., Boccardi, F., Dohler, M., *et al.*, 2014. Downlink and uplink decoupling: a disruptive architectural design for 5G networks. Proc. IEEE Globecom, p.1798-1803. http://dx.doi.org/10.1109/GLOCOM.2014.7037069

Feng, Z., Feng, Z., Li, W., *et al.*, 2014. Downlink and uplink splitting user association in two-tier heterogeneous cellular networks. Proc. IEEE Globecom, p.4659-4664. http://dx.doi.org/10.1109/GLOCOM.2014.7037543

Fooladivanda, D., Rosenberg, C., 2013. Joint resource allocation and user association for heterogeneous wireless cellular networks. *IEEE Trans. Wirel. Commun.*, **12**(1):248-257. http://dx.doi.org/10.1109/TWC.2012.121112.120018

Ge, X., Tu, S., Mao, G., *et al.*, 2016. 5G ultra-dense cellular networks. *IEEE Wirel. Commun.*, **23**(1):72-79. http://dx.doi.org/10.1109/MWC.2016.7422408

Jiang, L., Sha, X., Wu, X., *et al.*, 2016. A novel resource allocation algorithm based on downlink and uplink decouple access scheme in heterogeneous networks. *China Commun.*, **13**(6):1673-5447. http://dx.doi.org/10.1109/CC.2016.7513200

Jo, H.S., Sang, Y.J., Xia, P., *et al.*, 2012. Heterogeneous cellular networks with flexible cell association: a comprehensive downlink SINR analysis. *IEEE Trans. Wirel. Commun.*, **11**(10):3484-3495. http://dx.doi.org/10.1109/TWC.2012.081612.111361

Li, Y., Jin, D., Yuan, J., *et al.*, 2014. Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks. *IEEE Trans. Wirel. Commun.*, **13**(7):3965-3977. http://dx.doi.org/10.1109/TWC.2014.2325552

Lien, S., Chen, K., Lin, Y., 2011. Toward ubiquitous massive accesses in 3GPP machine-to-machine communications. *IEEE Commun. Mag.*, **49**(4):66-74. http://dx.doi.org/10.1109/MCOM.2011.5741148

López-Pérez, D., Güvenc, İ., Roche, G., *et al.*, 2011. Enhanced intercell interference coordination challenges in heterogeneous networks. *IEEE Wirel. Commun.*, **18**(3):22-30. http://dx.doi.org/10.1109/MWC.2011.5876497

Madan, R., Borran, J., Sampath, A., *et al.*, 2010. Cell association and interference coordination in heterogeneous LTE-A cellular networks. *IEEE J. Sel. Areas Commun.*, **28**(9):1479-1489. http://dx.doi.org/10.1109/JSAC.2010.101209

Nesterov, Y., Nemirovski, A., 1994. Interior point polynomial time methods in convex programming. *In*: SIAM Studies in Applied Mathematics. Philadelphia, PA.

Saad, W., Han, Z., Hjørungnes, A., 2011. Coalitional games for cooperative wireless cellular networks. *In*: Hossain, E., Kim, D., Bhargava, V. (Eds.), Cooperative Cellular Wireless Networks. Cambridge University Press.

Singh, S., Baccelli, F., Andrews, J.G., 2014. On association cells in random heterogeneous networks. *IEEE Wirel. Commun. Lett.*, **3**(1):70-73. http://dx.doi.org/10.1109/WCL.2013.111713.130707

Singh, S., Zhang, X., Andrews, J.G., 2015. Joint rate and SINR coverage analysis for decoupled uplink-downlink biased cell associations in HetNets. *IEEE Trans. Wirel. Commun.*, **14**(10):5360-5373. http://dx.doi.org/10.1109/TWC.2015.2437378

Smiljkovikj, K., Popovski, P., Gavrilovska, L., 2015. Analysis of the decoupled access for downlink and uplink in wireless heterogeneous networks. *IEEE Wirel. Commun. Lett.*, **4**(2):173-176. http://dx.doi.org/10.1109/LWC.2015.2388676

Son, K., Kim, H., Yi, Y., *et al.*, 2011. Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. *IEEE J. Sel. Areas Commun.*, **29**(8):1525-1536. http://dx.doi.org/10.1109/JSAC.2011.110903

Yang, C., Li, J., Guizani, M., 2016. Cooperation for spectral and energy efficiency in ultra-dense small cell networks. *IEEE Wirel. Commun.*, **23**(1):64-71. http://dx.doi.org/10.1109/MWC.2016.7422407

Ye, Q., Rong, B., Chen, Y., *et al.*, 2013. User association for load balancing in heterogeneous cellular networks. *IEEE Trans. Wirel. Commun.*, **12**(6):2706-2716. http://dx.doi.org/10.1109/TWC.2013.040413.120676