A Utility-Based Joint Resource Allocation Approach for Multi-Service in CoMP Networks

(article starts on next page)

A Utility-Based Joint Resource Allocation Approach for Multi-Service in CoMP Networks

Binru Huang · Jingya Li · Tommy Svensson

Abstract Coordinated multi-point (CoMP) joint transmission is considered as a key technique to mitigate inter-cell interference and improve the cell-edge performance in 3GPP LTE-Advanced. In this paper, a utility-based joint resource allocation approach is proposed to support CoMP joint transmission in orthogonal frequency division multiple access systems. Two different utility functions are considered with diverse quality of service provisioning. The objective is to maximize the sum utility of a clustered CoMP network, considering mixed real-time voice over IP (VoIP) and best-effort (BE) traffic patterns. A centralized joint resource allocation algorithm is proposed by decoupling the optimization problem into per-subchannel subproblems. Further more, a greedy user selection based low-complexity algorithm is proposed for practical implementation. For the sake of performance comparison, the proportional-fair based joint scheduling, as well as two non-coordinated scheduling schemes are also evaluated. Via system-level simulation we show that the proposed algorithms not only substantially improve the cell-edge spectrum efficiency of BE users, but also significantly decrease the packet drop ratio and call outage ratio of VoIP users.

B. Huang (✉) · J. Li · T. Svensson
Signals and Systems, Chalmers University of Technology,
41296 Gothenburg, Sweden
e-mail: binru@student.chalmers.se; aggiebinbin@gmail.com

J. Li
e-mail: jingya.li@chalmers.se

T. Svensson
e-mail: tommy.svensson@chalmers.se

1 Introduction

Future wireless networks are demanded to support multiple services, e.g., voice over IP (VoIP), video streams, file transfers, etc. To satisfy the diverse demands on data rate from multiple services, packet-based transmission is utilized to optimize the channel capacity in wireless networks. Meanwhile, orthogonal frequency division multiple access (OFDMA) is envisioned as a leading candidate to support broadband transmission and multimedia services [1]. Besides, the requirement for full frequency reuse is also imposed by wireless communications, in order to guarantee high spectrum efficiency and satisfy diverse quality of services (QoSs). However, inter-cell interference (ICI) due to the universal reuse of spectral resources can significantly degrade the system performance, and in turn impair the QoS in mixed-traffic wireless networks.

Coordinated Multi-Point (CoMP) joint transmission is proposed as a promising technique to improve cell-edge user throughput and system efficiency [1]. In CoMP systems, multiple coordinated cells are connected via a high-speed backbone. By using a joint transmission scheme in the downlink of CoMP systems, the ICI can be significantly mitigated by applying the signals transmitted from other cells to assist the transmission instead of acting as interference. In order to make the overhead of communication between coordinated base stations (BSs) affordable, clustering of BSs, i.e., dividing the network into small clusters of BS sectors (BSSs) [2–4], has been considered.

Clearly, radio resource management (RRM) cooperation among multiple cells plays a key role for controlling ICI, and in turn improving the system performance in CoMP networks. Currently, RRM schemes for CoMP systems are mainly studied for increasing cell-edge throughput and achieving the balance between system spectrum efficiency and user fairness, assuming all the users in the network have the same traffic modes [5–7]. In [5] and [6], algorithms utilizing coordinated scheduling and power control are proposed for controlling ICI. No joint transmission is taken into account. In [7], a utility-based algorithm for multi-cell coordinated resource allocation is proposed, supporting CoMP non-coherent joint transmission. Note that all these schemes treat the users equally without considering multiple traffic types in the CoMP networks.

Next generation wireless networks are challenged to meet the diverse QoS requirements imposed by various services [8]. The research on mixed-traffic scenarios is receiving more attention due to its significance in practical deployment of the Evolved UTRAN (E-UTRAN). Currently, research on resource allocation with mixed-traffic models focuses mainly on single-cell scenarios [8–10]. In [8], a unified approach based on utility functions to QoS-guaranteed scheduling is proposed for the downlink of time-division multiplexing (TDM). In [9], a mixed best-effort (BE) and VoIP traffic is studied, and a dynamic packet scheduling architecture is proposed to differentiate scheduling of different traffic classes. The result in [9] shows that with VoIP prioritizing the proposed algorithm keeps the VoIP UEs satisfied at the cost of decreased system spectral efficiency. In [10], a utility-based optimization is proposed for networks with mixed real-time and non-real-time traffic patterns, and it is shown to be able to satisfy the delay requirement of real-time traffic while balancing the fairness and efficiency. The main limitation of [8–10] is that RRM is designed only for the single cell scenario and no joint transmission is undertaken.

A utility-based scheduling and power control approach is proposed in [11] for CoMP networks with multiple services. However, only a flat fading channel model in the downlink transmission is considered. In this paper, we propose a packet-based joint resource allocation algorithm in OFDMA CoMP networks with mixed BE and VoIP services. We focus on the downlink of a CoMP network, with each BS having a fixed maximum transmit power

constraint. Two different utility functions are modeled for BE and VoIP users respectively. The objective is to maximize the sum utility of the users in each cluster. Binary power control (i.e., in any time slot, the cell either transmits with full power or does not transmit) is assumed. Assuming the power is equally allocated over multiple subchannels, hence, the analytical derivation shows that the optimization problem amounts to a user-group selection problem over all the subchannels, i.e., choosing user groups on all the subchannels to maximize the sum utility in each time slot. A centralized joint resource allocation algorithm is proposed by decoupling the optimization problem into per-subchannel sum utility maximizing problems. For practical implementation, a low-complexity distributed algorithm is also proposed based on the idea of greedy user selection (GUS) [12,13].

Both single-cluster and multi-cluster scenarios are investigated. Via the system level simulation, we show that our proposed algorithms well improve the performances at the cell-edge area in both of the two different scenarios. By taking advantage of the joint transmission scheme and utility differentiations, the presented algorithms significantly increase the cell-edge BE users efficiency, while suppressing the cell-edge packet drop ratio and call outage probability (i.e. the interruption probability of VoIP calls due to excessive packet drop ratio) for the VoIP service.

The rest of this paper is organized as follows. In Sect. 2, we provide the system model considered in this paper. In Sect. 3, the constrained optimization objective is formulated; a utility-based centralized joint resource allocation algorithm, and then a low-complexity distributed algorithm are proposed for radio resource allocation. Simulation results are presented in Sect. 4, for both single-cluster and multi-cluster scenarios. Finally, the conclusions are presented in Sect. 5.

2 System Model

2.1 Clustered CoMP Networks with Multi-Service

We focus on the downlink of a static clustered CoMP network with a set of BSSs (denoted by $\mathbb{N}$). The network is statically divided into several disjoint clusters, with each cluster consisting of three neighboring BSSs. A central unit (CU) is used to determine user scheduling and power control for all BSSs included in the cluster; see Fig. 1. According to the long term channel gain, users are divided into two classes, namely cell-center users (CCUs) and cell-edge users (CEUs), and we assume that joint transmission can only be applied to CEUs. Hence, we focus only on CEUs (denoted by $\mathbb{M}$) in this paper. The same spectrum bandwidth B is shared among all the BSSs, and equally divided into K subchannels. The BSSs are assumed to have one directional transmit antenna. Each BSS is assigned with a fixed maximum transmission power P, which is equally distributed over the K subchannels. Each CEU is equipped with one receive antenna and can only receive signals from a subset of the BSSs in the cluster where the CEU is located.

Each cluster is working independently in the network. Hence, without loss of generality, we consider a given cluster with a set $\mathcal{M}$ of CEUs and a set $\mathcal{N}$ of BSSs. In each time slot, the CU allocates users on all the subchannels for each BSS n based on the channel state information (CSI) of each CEU m. A user schedule index $x_{nm}^k(t)$ is defined as

$$x_{nm}^k(t) = \begin{cases} 1, & \text{BSS } n \text{ transmit to CEU } m \text{ on subchannel } k, \ n \in \mathcal{N}, \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$
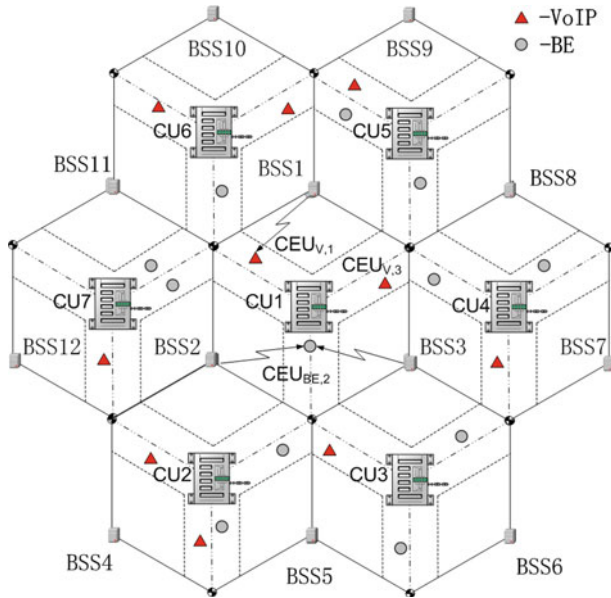
Fig. 1 System model for downlink joint transmission in clustered CoMP networks

at time slot t. Hence, the user schedule matrix can be denoted as $\mathbf{X}(t) = [x_{nm}^k(t)]$ with size $K \times N \times M$, where N and M are the cardinality of $\mathcal{N}$ and $\mathcal{M}$, respectively. Assume that each BSS can transmit to no more than one user in a time slot on each subchannel, and thus we have $\sum_{m=1}^M x_{nm}(t) \leq 1; \forall n \in \mathcal{N}$. The joint transmission BSS set for MS m on subchannel k can be denoted by $S_m^k = \{n | x_{nm}^k = 1, n \in \mathcal{N}\}$. Hence, according to (1), the interferring BSS set in the whole network for MS m on subchannel k is given by $\bar{S}_m^k = \{n' | \sum_{i \in \mathbb{M}, i \neq m} x_{n'i}^k = 1, n' \in \mathbb{N}\}$.

Let $P_n^k(t)$ denote the transmit power of BSS n on the subchannel k at time slot t. Assume binary power control on each subchannel, i.e., each BSS either transmits with the maximum power allocated on the subchannel $P_n^k(t) = P/K$, or does not transmit $P_n^k(t) = 0$. Let $G_{nm}^k(t)$ denote the channel gain between BSS n and CEU m on the subchannel k at time slot t, consisting of path-loss, and small-scale fading. Then with the power of the additive white Gaussian noise (AWGN) N_0, the signal to interference and noise ratio (SINR) of the CEU m on the subchannel k at time t based on non-coherent reception becomes

$$\gamma_m^k(t) = \frac{\sum_{n \in S_m^k} P/K \cdot G_{nm}^k(t)}{\sum_{j \in \bar{S}_m^k} P/K \cdot G_{jm}^k(t) + N_0}. \tag{2}$$

Hence, the achievable data rate of CEU m on the subchannel k using Shannon theorem is

$$R_m^k(t) = B \log_2 \left(1 + \beta \gamma_m^k(t)\right), \tag{3}$$

where β is related to the target bit error rate (BER), and given by [14] as $\beta = -1.5/\ln(5\text{BER})$. The instantaneous data transmission rate for CEU m at time slot t becomes

$$R_m(t) = \sum_{k=1}^K R_m^k(t). \tag{4}$$

2.2 Multiple Traffic Patterns

We consider two types of services in our system, i.e. BE and VoIP. The CEUs are further divided into two categories based on the services they require, i.e. BE users and VoIP users, denoted as CEU_{BE} and CEU_V, respectively in Fig. 1. The sets of BE and VoIP users are denoted as $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, with $\mathcal{M}_1 \cup \mathcal{M}_2 = \mathcal{M}$ and $\mathcal{M}_1 \cap \mathcal{M}_2 = \varnothing$. According to the different characteristics of the two services, two different types of traffic models are considered. Full buffer is assumed for BE users. For the VoIP service, which has a strict requirement on latency, we assume a very bursty traffic with a low bit rate requirement. A maximum allowed instantaneous queuing delay $\tau_{\max}$ is imposed for QoS, i.e., $\tau_m(t) \leq \tau_{\max}$, where $\tau_m(t)$ denotes the instantaneous packet delay of VoIP user m. The expired packets with excessive latency larger than $\tau_{\max}$ are discarded at the transmitter side. According to the different service requirements for each type of users, two different utility functions are defined to represent their satisfaction. For a BE user, the satisfaction is assumed to depend on its average throughput. Hence, the utility function of BE user m is defined as a monotonically increasing function of the average throughput $\bar{R}_m(t)$ at time t

$$U_m(t) = U_{BE,m}\left(\bar{R}_m(t)\right); \forall m \in \mathcal{M}_1, \tag{5}$$

where $\bar{R}_m(t)$ is estimated using an exponential filter as [15]

$$\bar{R}_m(t) = (1 - \rho_{BE})\bar{R}_m(t-1) + \rho_{BE}R_m(t). \tag{6}$$

The satisfaction of a VoIP user is assumed to depend on its service delay. Hence, the utility function for a VoIP user is defined as a monotonically decreasing function of its average queuing delay, which is given by

$$U_m(t) = U_{V,m}\left(\bar{d}_m(t)\right); \forall m \in \mathcal{M}_2, \tag{7}$$

where $\bar{d}_m(t)$ is its average queuing delay at time t. We estimate $\bar{d}_m(t)$ through the approach proposed in [16]. Define $Q_m(t)$ as the queue size in bits at the end of time slot t, and $\alpha_m(t)$ as the instantaneous arriving bits at the end of slot t. With departure rate $R_m(t)$, the queue size is calculated by

$$Q_m(t) = Q_m(t-1) - \min\{R_m(t)T_s, Q_m(t)\} + \alpha_m(t-1), \tag{8}$$

where T_s is the slot duration. Assuming that $Q_m(t)$ is ergodic, then with Littles Law the average delay can be estimated by $\bar{d}_m(t) = \bar{Q}_m(t)/\bar{\alpha}_m$, where $\bar{\alpha}_m$ denotes the time averaging arrival bits per slot, and $\bar{Q}_m(t)$ is the average queue size at time slot t. Similar to $\bar{R}_m(t)$ in (6), the average queue size is estimated by

$$\bar{Q}_m(t) = (1 - \rho_V)\bar{Q}_m(t-1) - \rho_V Q_{V,m}(t). \tag{9}$$

This in turn leads to the estimate of the average delay via

$$\bar{d}_m(t) = (1 - \rho_V)\bar{d}_m(t-1) + \rho_V\bar{\alpha}_m^{-1}Q_m(t). \tag{10}$$

Substitute (9) into (10), and the estimate of the average delay is ultimately expressed by

$$\begin{aligned}
\bar{d}_m(t) = (1 - \rho_V)\bar{d}_m(t-1) &+ \rho_V\bar{\alpha}_m^{-1}[Q_m(t-1) \\
&- \min\{R_m(t)T_s, Q_m(t-1)\} + \alpha_m(t-1)].
\end{aligned} \tag{11}$$

3 Utility-Based Joint Resource Allocation

In this section, the maximum sum utility optimization problem is formulated. Then, a utility-based joint resource allocation algorithm, as well as a low-complexity algorithm, are proposed for independent radio resource allocation in each CoMP cluster.

3.1 Problem Formulation

Our objective is to maximize the sum utility of the CEUs within a CoMP cluster. The objective function is formulated as

$$U(t) = \sum_{m_1 \in \mathcal{M}_1} U_{BE,m_1} \left(\bar{R}_{m_1}(t) \right) + \sum_{m_2 \in \mathcal{M}_2} U_{V,m_2} \left(\bar{d}_{m_2}(t) \right). \tag{12}$$

Note that $\bar{R}_{m_1}(t-1)$ and $\bar{d}_{m_2}(t-1)$ are known at time slot t. Hence, using Taylor expansion, to maximize (12) at time slot t corresponds to maximize [16]

$$\Pi(t) = \sum_{m_1 \in \mathcal{M}_1} U'_{BE,m_1} \left(\bar{R}_{m_1}(t-1) \right) \bar{R}_{m_1}(t) $$
$$+ \sum_{m_2 \in \mathcal{M}_2} U'_{V,m_2} \left(\bar{d}_{m_2}(t-1) \right) \bar{d}_{m_2}(t). \tag{13}$$

Substitute (6) and (11) into (13), and let the CU control the service bit rate so that $R_{m_2} T_s \leq Q_{m_2}(t-1)$. Then with fixed $Q_{m_2}(t-1)$ and $\alpha_{m_2}(t-1)$ at slot t, $\Pi(t)$ can be reformulated as a function only of $R_m(t)$ as

$$\Pi(t) = \sum_{m_1 \in \mathcal{M}_1} \rho_{BE} U'_{BE,m_1} \left(\bar{R}_{m_1}(t-1) \right) R_{m_1}(t) $$
$$- \sum_{m_2 \in \mathcal{M}_2} \rho_V \bar{\alpha}_{m_2}^{-1} T_s U'_{V,m_2} \left(\bar{d}_{m_2}(t-1) \right) R_{m_2}(t). \tag{14}$$

Note that the marginal utility functions $U'_{BE,m_1}(\cdot)$ and $U'_{V,m_2}(\cdot)$ are related to the scheduling weights or priorities, and thus play a key role in the scheduler design. Since $R_m(t)$ is related to $\mathbf{X}(t)$, $\Pi(t)$ turns out to be a function of $\mathbf{X}(t)$. Using $\Pi(\mathbf{X}(t))$ to represent $\Pi(t)$, (14) becomes

$$\Pi(\mathbf{X}(t)) = \sum_{m_1 \in \mathcal{M}_1} \pi_{m_1} R_{m_1}(\mathbf{X}(t)) + \sum_{m_2 \in \mathcal{M}_2} \pi_{m_2} R_{m_2}(\mathbf{X}(t)), \tag{15}$$

where π_{m_1} and π_{m_2} are fixed at time t, with

$$\begin{cases} \pi_{m_1} = U'_{BE,m_1} \left(\bar{R}_{m_1}(t-1) \right) \rho_{BE}, \\ \pi_{m_2} = -U'_{V,m_2} \left(\bar{d}_{m_2}(t-1) \right) \rho_V \bar{\alpha}_{m_2}^{-1} T_s. \end{cases} \tag{16}$$

Ultimately, the optimization problem for each cluster is mathematically formulated as

$$\max_{\mathbf{X}(t)} \Pi(\mathbf{X}(t))$$
$$\text{s.t. } 1) R_{m_2} T_s \leq Q_{m_2}(t-1), \forall m_2 \in \mathcal{M}_2$$
$$2) \sum_{m=1}^{M} x_{nm}(t) \leq 1, \forall n \in \mathcal{N}$$
$$3) \tau_{m_2}(t) \leq \tau_{\max}, \forall m_2 \in \mathcal{M}_2. \tag{17}$$

That is, the optimization problem (17) turns out to be to find the $\mathbf{X}^*(t)$ that maximizes $\Pi(\mathbf{X}(t))$ in (15) under the constraints (1) the instantaneous departure data size for VoIP users can be no more than the attainable waiting queue size, (2) a BSS transmits to at most

one CEU, and (3) the instantaneous delay of a VoIP packet should be no greater than the maximum allowed delay.

3.2 Algorithm Description

Based on the assumption of equal power allocation over subchannels, the optimization problem in (17) is decoupled into maximizing sum utility independently on each subchannel. Thus, a utility-based centralized joint resource allocation algorithm is proposed in this section for the OFDMA CoMP system. In order to further reduce the computational complexity, inspired by the GUS [12,13] based scheme, a low-complexity distributed algorithm is also proposed for practical implementation.

3.2.1 Utility-Based Centralized Joint Resource Allocation Algorithm

Consider a given cluster with N BSSs, M users, and K subchannels. Let $\mathbb{X}(t)$ denote the set of all feasible user schedules per subchannel in each time slot. With binary power control, we have $|\mathbb{X}(t)| = (M + 1)^N$, where $|\mathbb{X}(t)|$ denotes the cardinality of the set $\mathbb{X}(t)$. Hence, with K subchannels in total, the complexity for finding the optimal user groups over all the subchannels is $O\left(K \cdot (M + 1)^N\right)$.

At each time slot t, as $\mathbf{X}_k(t)$ denotes the user schedule matrix on the subchannel k, and $\Pi(\mathbf{X}_k(t)) = \sum_{m \in \mathcal{M}} \pi_m R_m^k(\mathbf{X}_k(t))$ represents the equivalent sum utility on the subchannel k. On each subchannel k, the algorithm starts with en empty user set and the transmit power is initialized as P/K. An exhaustive search is done in the set $\mathbb{X}(t)$ of all the feasible user schedules for the optimal user group $\mathbf{X}_k^*(t)$ that gives the largest $\Pi(\mathbf{X}_k(t))$. After centralized joint user selection on each subchannel, the queue size of each VoIP user is adjusted such that the spectrum is not wasted serving empty queues. Hence, the estimate of the queue size of VoIP user m on the subchannel k is updated as

$$\hat{Q}_m^{(k)} = \hat{Q}_m^{(k-1)} - R_m^k T_s \sum_{n=1}^{N} x_{m,n}^k, \ m \in \mathcal{M}_2. \tag{18}$$

At the end of each time slot t, $\bar{R}_m(t)$, $\bar{d}_m(t)$, and $\bar{Q}_m(t)$ are updated based on $\mathbf{X}_k^*(t)$. The algorithm is outlined in Algorithm 1.

Algorithm 1 Utility-Based Centralized Joint Resource Allocation Algorithm

---

1: Initialize $P_n^k = P/K$, $\forall n \in \mathcal{N}$; $\hat{Q}_m^{(1)} = Q_m(t-1)$, $m \in \mathcal{M}_2$; $R_m(0) = 0$, $Q_m(0) = 0$, $\forall m \in \mathcal{M}$, $\forall k \in \mathcal{K} = \{1, 2, ..., K\}$
2: **for** each time slot t **do**
3:   **for** $k = 1$ to K **do**
4:     **for all** schedules $\mathbf{X}_k(t)$ in $\mathbb{X}(t)$ **do**
5:       Compute $\Pi(\mathbf{X}_k(t))$ using (15)
6:     **end for**
7:     $\mathbf{X}_k^*(t) = \arg\max_{\mathbf{X}_k(t) \in \mathbb{X}(t)} \Pi(\mathbf{X}_k(t))$
8:     Update $\hat{Q}_{m_2}^{(k)}$, $\forall m_2 \in \mathcal{M}_2$ using (18)
9:   **end for**
10:   Update $\bar{R}_{m_1}(t)$, $\forall m_1 \in \mathcal{M}_1$ using (6)
11:   Update $\bar{d}_{m_2}(t)$, $\bar{Q}_{m_2}(t)$, $\forall m_2 \in \mathcal{M}_2$ using (11), (9)
12:   $t = t + 1$
13: **end for**

3.2.2 Utility-Based Distributed Joint Resource Allocation Algorithm

Inspired by GUS, a low-complexity distributed algorithm is proposed for joint resource allocation for the OFDMA CoMP system. In this algorithm, we apply GUS combined with binary power control on each subchannel. We starts from an empty user decision set S_k on the subchannel k for the group of coordinated BSSs in a CoMP cluster. A user is pre-selected from all the available users, if it can achieve the highest sum utility-based on the current scheduled user set S_k. Furthermore, the pre-selected user is added into the scheduled user set S_k only if the total utility will be increased. The GUS combined with transmit power allocation is performed successively for each BSS in the cluster. Then, the subchannel is assigned to the coordinated BSSs based on the scheduled user set S_k in the same way as the second step of the proposed centralized joint resource allocation algorithm. Hence, the computational complexity of this algorithm is reduced to $O(K \cdot N \cdot M)$, achieving a substantial reduction by contrast to the centrailized algorithm. The algorithm is described in Algorithm 2.

Algorithm 2 Utility-Based Distributed Joint Resource Allocation Algorithm

---

1: Initialize $P_n^k = P/K, \forall n \in \mathcal{N}$; $\hat{Q}_m^{(1)} = Q_m(t-1), m \in \mathcal{M}_2$; $R_m(0) = 0, Q_m(0) = 0, \forall m \in \mathcal{M}, \forall k \in \mathcal{K} = \{1, 2, ..., K\}$

2: **for** each time slot t **do**

3:    Initialize $S_k = \varnothing, \mathbf{X}_k(t) = \mathbf{0}$

4:    **for** $k = 1$ to K **do**

5:      **for** $n = 1$ to N **do**

6:        $m^* = \arg\max_{m \in \mathcal{M}} \Pi \left(\mathbf{X}_k(t)|_{x_{nm}^k = 1} \right)$

7:        **if** $\Pi \left(\mathbf{X}_k(t)|_{x_{nm^*}^k = 1} \right) > \Pi \left(\mathbf{X}_k(t)|_{x_{nm^*}^k = 0} \right)$ **then**

8:          $S_k = S_k \cup m^*, \mathbf{X}_k(t)|_{nm^*} = 1$

9:        **end if**

10:     **end for**

11:    Update $\hat{Q}_{m_2}^{(k)}, \forall m_2 \in \mathcal{M}_2$ using (18)

12:    **end for**

13:    Update $\bar{R}_{m_1}(t), \forall m_1 \in \mathcal{M}_1$ using (6)

14:    Update $\bar{d}_{m_2}(t), \bar{Q}_{m_2}(t), \forall m_2 \in \mathcal{M}_2$ using (11), (9)

15:    $t = t + 1$

16: **end for**

---

4 Simulation Results

We focus on the downlink of an OFDMA cellular system. The total bandwidth is 3 MHz, which is divided into 15 parallel subchannels. The carrier frequency is 2 GHz. We randomly drop the two different types of users in the CoMP network. The cell radius is set to 500 m. Define the path loss with $L(d) = 128.1 + 37.6 \log_{10} d$ in dB, where d is the distance in km [17], where the long term channel gain is under the threshold -100 dB. 1,000 independent trials are evaluated by Monte-Carlo simulation under various numbers of CEUs per cluster.

A number of BE and VoIP users are uniformly allocated in the cell-edge area of the CoMP cluster, respectively, with the probability of 50 % for each. The target BER for data transmission is prescribed as 10^{-5}. The utility function for BE users is defined as

$$U_{BE,m_1} \left(\bar{R}_{m_1}(t) \right) = \ln \left(\bar{R}_{m_1}(t) \right). \tag{19}$$

For the VoIP users we consider a VoIP traffic model with packet inter-arrival time of 10 ms, and packet size of 40 bytes. Regarding QoS requirements, we set the maximum queuing delay to be 20 ms. Packets with packet delays greater than the maximum delay $\tau_{\max}$ or unsuccessfully received at the reciver side due to the poor channel condition are discarded. VoIP calls for users experiencing packet drop ratio higher than 2 % result in call outage. Hence, the utility function is defined as

$$U_{V,m_2}\left(\bar{d}_{m_2}(t)\right) = -\frac{\log_{10}\delta_{m_2}}{2\check{d}_{m_2}}\left(\check{d}_{m_2}^2 - \bar{d}_{m_2}^2(t)\right), \qquad 0 < \delta_{m_2} < 1, \tag{20}$$

where $\check{d}_{m_2}$ is the maximum allowable queuing delay for VoIP users; $\delta_{m_2} = 0.1$ is a constant, chosen to balance the priorities of different types of users [18]. Recall (14), the BE users with lower average throughput can get higher priority in the scheduling if $U_{BE,m_1}(\cdot)$ is chosen as in (19). Similarly, with $U_{V,m_2}(\cdot)$ defined as in (20), the VoIP users will gain higher priorities if they experience larger delays. In fact, the marginal utility function of (20) turns out to be the largest-weighted-average-delay-first (LWADF) scheduling [18], i.e., users in the queue experiencing the largest average delay have the highest priorities and should be served first in each round of scheduling. Besides, ρ_{BE} and ρ_V also play an important role in balancing priorities of the two types of users in scheduling, and are prescribed as $\rho_{BE} = 0.01$ and $\rho_V = 0.05$, respectively.

The average sum utility is evaluated as the assessment of the proposed utility-based centralized joint resource allocation algorithm (C-CUBP), as well as the proposed utility-based low-complexity distributed joint resource allocation algorithm (C-DUBP). Meanwhile, as a special case of the two aforementioned schemes, algorithms with the power of all BSSs in the cluster always turned on, named C-CUB and C-DUB respectively, are also considered and assessed. Furthermore, three other algorithms are considered for comparison:

1. *Coordinated proportional-fair scheduling without power control (C-PF)*: The algorithm is aimed to maximize the proportional throughput-fair index [7] with joint transmission, but the differentiations of diverse traffic models are not considered.
2. *Utility-based scheduling without joint transmission or power control (NC-UB)*: Similar to the proposed C-CUB algorithm, but no joint transmission is supported.
3. *Proportional-fair scheduling without joint transmission or power control (NC-PF)*: Similar to C-PF but no joint transmission is supported either.

Besides the average sum utility, the average user throughput for the cell-edge BE users, and the average packet drop ratio and call outage ratio for the cell-edge VoIP traffic users, respectively, are also investigated.

4.1 Single-Cluster Performance Analysis

Firstly, we focus on the single-cluster CoMP scenario, where a cluster of three BSSs is considered. Hence, only intra-cluster interference is taken into account. The true SINR expression for user m on the subchannel k is then given as

$$\gamma_m^k = \frac{\sum_{n\in\mathcal{N}} P/K \cdot G_{nm}^k x_{nm}^k}{\sum_{i\in\mathcal{N}} P/K \cdot G_{im}^k \left(\sum_{\substack{s\in\mathcal{M}\\s\neq m}} x_{is}^k\right) + N_0}. \tag{21}$$

We investigate different cases with different number of CEUs for the aforementioned algorithms. In Fig. 2, the cell-edge average sum utility of the algorithms is plotted with respect to the number of CEUs. It can be seen that the proposed centralized joint resource allocation
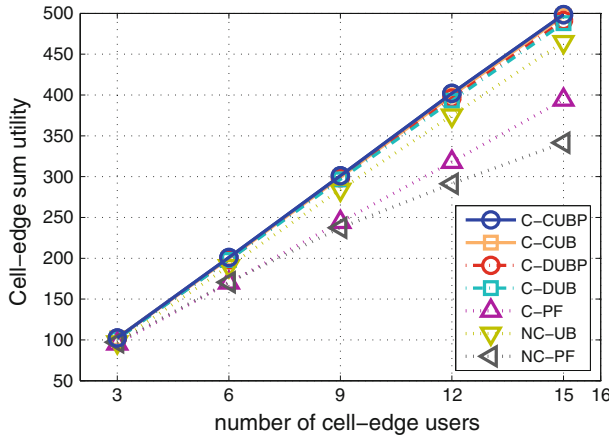
Fig. 2 Cell-edge sum utility of the single-cluster CoMP system versus number of cell-edge users

algorithm C-CUBP achieves the highest sum-utility performance. The C-CUB algorithm without power control still achieves proximate performance, with slightly lower sum utility than that of the C-CUBP algorithm. The proposed distributed algorithms C-DUBP and C-DUB also provide an appropriate performance with more than 99 and 98 %, respectively, of the sum utility from the centralized C-CUBP algorithm, but with much lower complexity. By exploiting the two different utility functions, even the NC-UB algorithm without joint transmission outperforms the two proportional-fair scheduling based algorithms C-PF and NC-PF, which do not take the differentiations of diverse traffic models into account.

Moreover, with binary power control, a better power saving is provided by the C-CUBP algorithm, with the average ratio of the BSSs turned on as 66.82–70.93 %, compared to 100 % for C-CUB. However, the simulation results show that the average power-on ratio of C-DUBP is 96.5–97.4 %, which provides no better power saving compared to the C-CUBP algorithm.

To further improve our understanding of the proposed algorithms, we then evaluate the performances of BE users and VoIP users, separately. We can see from Fig. 3 that as the
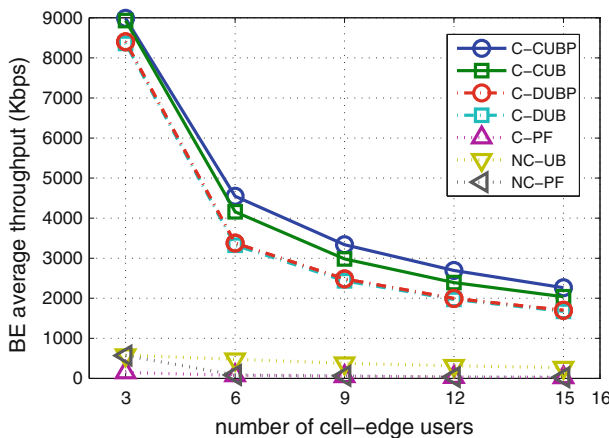

Fig. 3 Cell-edge BE average throughput of the single-cluster CoMP system versus number of cell-edge users

traffic gets heavier, the average throughput of the BE users decreases for all the algorithms. However, the proposed centralized C-CUBP achieves the highest average throughput of BE users in all cases. The C-CUB algorithm still achieves 92 % of the BE average throughput from the centralized C-CUBP algorithm, even without binary power control. The two low-complexity algorithms C-DUBP and C-DUB yield lower BE average throughput, achieving about 78 and 77 % from that of the C-CUBP, respectively. The forementioned four algorithms all significantly outperform the NC-UB algorithm, in which no joint transmission is considered. Hence, joint transmission contributes to the substantial improvement in the average throughput of the BE users. Moreover, by considering different utility functions, the NC-UB algorithm still yields much better performance than the C-PF and NC-PF algorithms.

In Figs. 4 and 5, we plot the average packet drop ratio and call outage ratio for cell-edge VoIP users, respectively. From Fig. 4, it can be seen that the algorithms considering joint transmission (i.e., C-CUBP, C-CUB, C-DUBP, C-DUB and C-PF) all outperform those without considering joint transmission. The packet drop ratios of those algorithms with joint transmission are all less than 0.01 %. Among them the centralized proportional-fair
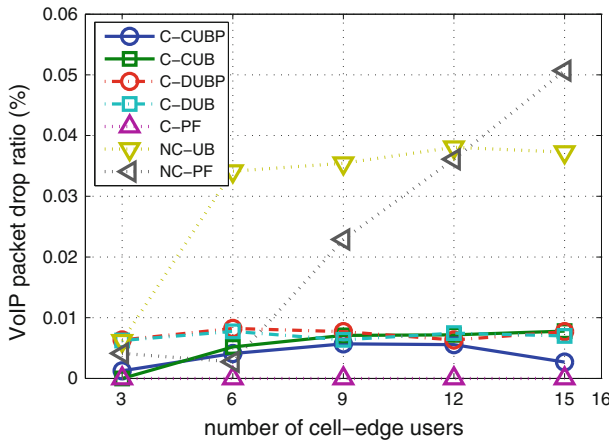

Fig. 4 Cell-edge VoIP packet drop ratio of the single-cluster CoMP system versus number of cell-edge users
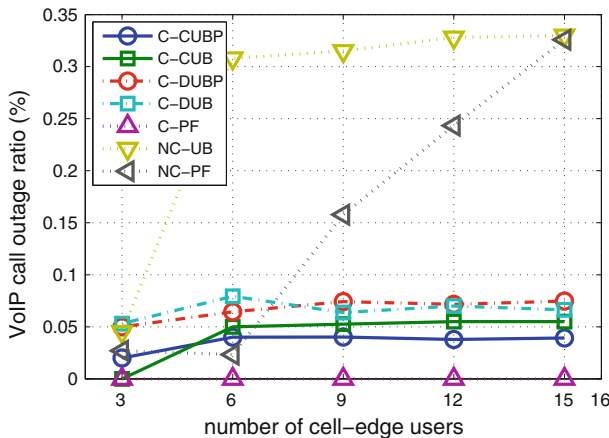

Fig. 5 Cell-edge VoIP call outage ratio of the single-cluster CoMP system versus number of cell-edge users

scheduling based C-PF algorithm achieves the best performance for VoIP users, however, at the cost of the impaired performance of BE users according to the plottings in Fig. 3. In contrast, our proposed utility-based algorithms (i.e., C-CUBP, C-CUB, C-DUBP, and C-DUB) can also achieve low packet drop ratios for VoIP users, and meanwhile provide an appropriate performance for BE users. In Fig. 5, we can observe that the call outage ratios are lower than 0.1 % for all the proposed utility-based joint transmission algorithms.

Hence, we can conclude that the improved cell-edge performances for both BE and VoIP users in the single-cluster CoMP scenario, are yielded by taking advantage of joint transmission. Moreover, by exploiting good diverse QoS provisioning through exploiting utility functions, the utility-based algorithms achieve a better balance between improving the average throughput of cell-edge BE users, and satisfying QoS requirements for cell-edge VoIP users, than the proportional fair algorithms. Meanwhile, binary power control could offer a better power saving.

4.2 Multi-Cluster Performance Analysis

To evaluate the impact of the inter-cluster interference, we investigate the performances of all the algorithms in a multi-cluster network in this subsection. We consider multiple static disjoint clusters, with each cluster consisting of three adjacent BSSs. Accordingly, the estimated SINR expression of user m on the subchannel k turns out to be

$$\gamma_m^k = \frac{\sum_{n \in \mathcal{N}} P/K \cdot G_{nm}^k x_{nm}^k}{\sum_{i \in \mathcal{N}} P/K \cdot G_{im}^k \left(\sum_{\substack{s \in \mathcal{M} \\ s \neq m}} x_{is}^k \right) + I_m^k + N_0}, \tag{22}$$

where $I_m^k = \sum_{j \notin \mathcal{N}} P/K \cdot G_{jm}^k$ is the estimated cochannel ICI on subchannel k for MS m from neighboring clusters, assuming that all the BSSs in the neighboring clusters are turned on.

The simulation results show that by considering the inter-cluster interference the cell-edge sum utility of all the considered different algorithms are degraded. From Fig. 6, we can see that the proposed centralized algorithm C-CUBP algorithm still achieves the highest sum-utility performance in the multi-cluster scenario, approximately 96 % of the sum utility of the single-cluster scenario. The performance of the proposed C-CUB, C-DUBP, and C-DUB algorithms approaches approximately 96.3, 92.1, and 92.5 % of the sum utility of the single-cluster scenario, respectively. The C-PF algorithm approaches more than 98 % of the sum utility of the single-cluster scenario, but this is still lower than the sum utility of our proposed algorithms. That is because the C-PF algorithms concentrates only on the spectrum efficiency fairness for all the users, but VoIP traffic has low constant average data rate, so for either the single-cluster or the multi-cluster scenario, the C-PF algorithm will yield similar performances.

By contrast to the aforementioned algorithms with joint transmission, the sum utility of the NC-UB algorithm is significantly degraded. It achieves only about 70 % of that of the single-cluster scenario, even much lower than the C-PF algorithm. The NC-PF algorithm yields even much worse performance than in the single-cluster scenario. Hence, we can conclude that inter-cluster interference has more significant impairment on the algorithms without considering joint transmission, than on those considering joint transmission. Therefore, the improved sum utility of the proposed algorithms, compared to the other algorithms, is yielded by taking advantage of joint transmission.
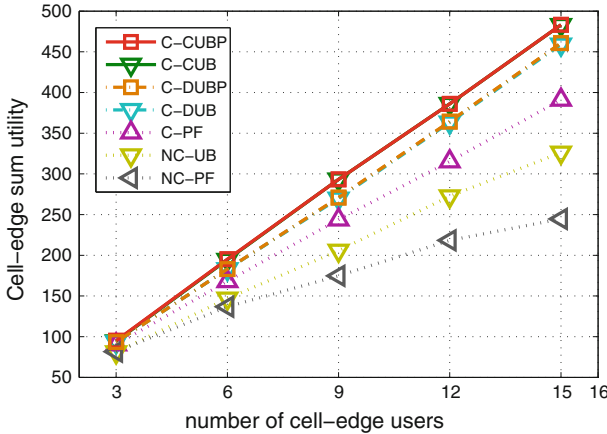
Fig. 6 Cell-edge sum utility of the multi-cluster CoMP system versus number of cell-edge users
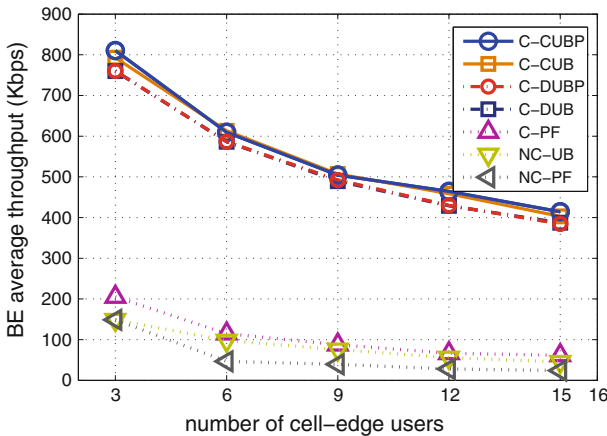

Fig. 7 Cell-edge BE average throughput of the multi-cluster CoMP system versus number of cell-edge users

Moreover, the power-on ratio for C-CUBP and C-DUBP is 75.94–88.6 % and 75.95–85.67 %, respectively. Hence, the proposed low-complexity power control algorithm C-DUBP achieves a better power saving even though it is degraded in the sum utility in the multi-cluster scenario.

The variations of the sum utility in the multi-cluster scenario can be explained more clearly by Figs. 7, 8 and 9. Compared to the BE average throughput of the single-cluster scenario as in Fig. 3, a substantial decrease in the BE average throughput could be observed in Fig. 7 for all the considered algorithms except the C-PF algorithm, due to the inter-cluster interference. However, our proposed utility-based centralized joint transmission and power control algorithm C-CUBP still achieves the highest BE average throughput. The proposed centralized algorithm C-CUB and the two low-complexity algorithms also outperform the three other comparison algorithms. The C-PF algorithm yields proximate but still low BE average throughput, compared to that in the single-cluster scenario as in Fig. 3.

Compared with the VoIP performances of single-cluster scenario as in Figs. 4 and 5, we can see from Figs. 8 and 9 that there is also impairment on the performance of the VoIP users in
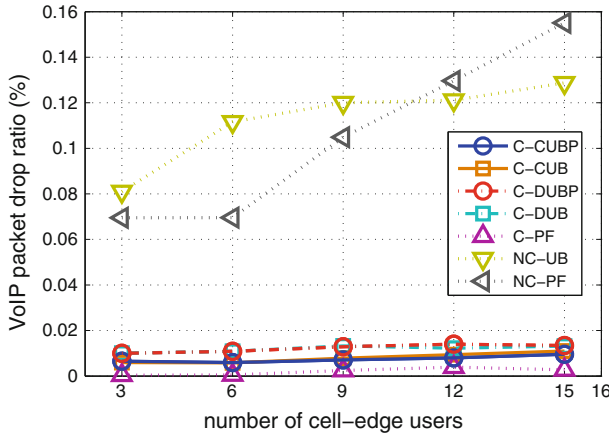
Fig. 8 Cell-edge VoIP packet drop ratio of multi-cluster CoMP system versus number of cell-edge users
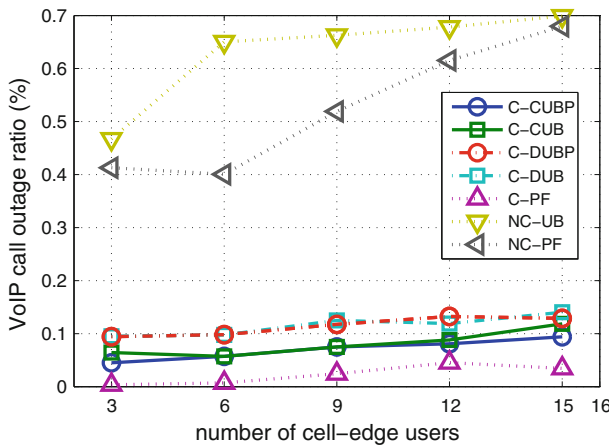

Fig. 9 Cell-edge VoIP call outage ratio of multi-cluster CoMP system versus number of cell-edge users

the multi-cluster scenario. With inter-cluster interference, both the average packet drop ratio and call outage ratio increases as the traffic gets heavier for all the algorithms. However, the algorithms with joint transmission still achieve a good performance for VoIP services. Without considering joint transmission, NC-UB and NC-PF achieve even worse performance for the VoIP services than in the single-cluster scenario, with substantial increase in the packet drop ratio and call outage ratio degradation in the multi-cluster scenario. Hence, joint transmission contributes to suppressing the impairments on the VoIP performances caused by inter-cluster interference.

In general, the performance of all the considered algorithms degrades at the cell-edge area for both BE and VoIP users, resulting from inter-cluster interference in the multi-cluster scenario. Especially for BE users, the performance is degraded significantly in terms of the cell-edge average throughput, even for the C-CUBP algorithm that achieves the best performance. However, with joint transmission our proposed algorithms could still provide an appropriate average throughput for BE users, and meanwhile guarantee the QoSs for cell-edge VoIP users.

5 Conclusions

In this paper, we consider the downlink of clustered CoMP networks, assuming that joint resource allocation is performed independently within each cluster. A utility-based joint resource allocation approach is proposed to maximize the cluster sum utility, considering mixed BE and VoIP traffic. Firstly, a centralized algorithm is developed to jointly assign a group of users on each subchannel. Targeting at practical scenarios, a low-complexity distributed algorithm based on GUS is then proposed. We evaluate the performances of the proposed algorithms in both a single-cluster and a multi-cluster scenario. Via system level simulation results we show that the proposed algorithms can provide a great improvement in the average throughput of BE users, and meanwhile substantially suppress the average packet drop ratio and call outage ratio of VoIP users. It can also be seen that the performance of the proposed algorithms are degraded in the multi-cluster scenario, due to the inter-cluster interference. However, an appropriate throughput for BE users as well as the QoSs for VoIP users can still be guaranteed by exploiting joint transmission and utility diversities. Moreover, binary power control could also provide a better power saving compared to those algorithms without power control.

References

1. 3GPP TR 36.814, V9.0.0. (2010). *Further advancements for EUTRA physical layer aspects*.
2. Zhang, J., Chen, R., Andrews, J. G., Ghosh, A., & Heath, R. W. (2009). Networked MIMO with clustered linear precoding. *IEEE Transactions Wireless Communications*, 1910–1921.
3. Boccardi, F., & Huang, H. (2007). Limited downlink network coordination in cellular networks. *In Proceeding PIMRC'07*.
4. Li, J., Botella, C., & Svensson, T. (2012). Resource allocation for clustered network MIMO OFDMA systems. *EURASIP Journal on Wireless Communications and Networking.*.
5. Venturino, L., Prasad, N., & Wang, X. (2009). Coordinated scheduling and power allocation in downlink multicell OFDMA networks. *IEEE Transactions Vehiculat Technology*, *58*, 2835–2848.
6. Chen, L., Chen, W., Zhang, X., Zhang, Y., Xiong, C., & Yang, D. (2009). Intercell coordinated resource allocation for mobile WiMAX system. In *Proceedings IEEE WCNC* (pp. 1–6).
7. Li, J., Xu, X., Chen, X., Tao, X., Zhang, H., Svensson, T., et al. (2010). Downlink radio resource allocation for coordinated cellular OFDMA networks. *IEICE Transactions on Communications*, *a93–b*, 3480–3488.
8. Wang, X., Giannakis, G. B., & Marques, A. G. (2007). A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks. *Proceedings of the IEEE*, *95*, 2410–2431.
9. Puttonen, J., Kolehmainen, N., Henttonen, T., Moisio, M. & Rinne, M. (2008). Mixed traffic packet scheduling in UTRAN Long Term Evolution Downlink. In *Proc. PIMRC'08* (pp. 1–5).
10. Katoozian, M., Navaie, K. & Yanikomeroglu, H. (2009). Utility-based adaptive radio resource allocation in OFDM wireless networks with traffic prioritization. *IEEE Transactions Wireless Communications, 8*.
11. Huang, B., Li, J., & Svensson, T. (2011). A utility-based scheduling approach for multiple services in coordinated multi-point networks. *IEEE Wireless Personal Multimedia Communications*.
12. Skjevling, H., Gesbert, D., & Hjørungnes, A. (2008). Low-complexity distributed multibase transmission and scheduling. *EURASIP Journal Advanced Signal*.
13. Li, J., Svensson, T., Botella, C., Eriksson, T., Xu, X., & Chen, X. (2011). Joint scheduling and power control in coordinated multi-point clusters. *VTC Fall*, 1–5.
14. Goldsmith, A. J., & Chua, S.-G. (1997). Variable-rate variable-power MQAM for fading channels. *IEEE Transactions Communications*, *45*, 1218–1230.
15. Song, G., & Li, Y. (2005). Cross-layer optimization for OFDM wireless networks-part II: Algorithm development. *IEEE Transactions Wireless Communications*, *4*, 625–634.

16. Song, G., & Li, Y. (2005). *Cross-layer resource allocation and scheduling in wireless multicarrier networks, Ph.D. dissertation*. :Georgia Institution Technology .
17. 3GPP TR 25.814 v7.1.0. (2006). *Physical layer aspects for evolved UTRA (Release 7)*.
18. Andrews, M., Kumaran, K., Ramanan, K., Stolyar, A. L., Vijayakumar, R., & Whiting, P. (2001). Providing quality of service over a shared wireless link. *IEEE Communications Magazine*, *39*, 150–154.

Author Biographies

Binru Huang received her B.E. in Electronic Information Engineering in Beijing University of Aeronautics and Astronautics (BUAA), Beijing, China, in 2006. From Sep. 2009 to Oct. 2011, she was a research master student in the master program of Communication Engineering at the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. Since Oct. 2011 till now she has been working at Ericsson AB, Sweden. Her research interests are in the broad areas of cooperative wireless communications, with a focus on resource allocation and interference management techniques for multi-cell multi-user network systems.

Jingya Li received her B.E. in Communication Engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2008. From Sept. 2008 to Sept. 2010, she was a research master student at the Department of Communication and Information Systems, BUPT, Beijing, China. Since Oct. 2010, she has been pursuing her Ph.D. at the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. Her research interests are in the broad areas of cooperative wireless communications, with a focus on resource allocation and interference management techniques for multi-cell multi-user network MIMO systems.

Tommy Svensson is Associate Professor in Communication Systems at Chalmers University of Technology, where he is leading the research on air interface and microwave backhauling technologies for future wireless systems. He has a PhD in Information theory from Chalmers in 2003, and he has been working at Ericsson AB with core networks, radio access networks and microwave transmission products. His main research interests are in design and analysis of physical layer algorithms, multiple access, resource allocation, cooperative systems and moving relays. He is responsible for the Master's program on Communication Engineering at Chalmers, senior member of IEEE, and chairman of the Swedish IEEE VT/COM/IT chapter.