This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

Training Resource Allocation for User-centric Base Station Cooperation Networks

Zhilin Chen, Xueying Hou, and Chenyang Yang

***Abstract*—User-centric base stations (BSs) cooperative transmission strives to satisfy the quality of service of each user no matter where the user is located. The resulting user-dependent cooperative clusters are inevitably overlapped. To minimize mean square error of channel estimation assisting user-centric downlink cooperative transmission, the training signals sent from the BSs in each cluster or from the users selecting the same BS in their clusters should be mutually orthogonal. In this paper, we study orthogonal training resource allocation problem for user-centric cooperative network aiming at minimizing the overall training overhead. We find the optimal solution through a graph-theoretic approach. To provide a feasible solution for large scale networks, a low complexity algorithm is then proposed. Simulation results show that the algorithm performs closely to the optimal solution, and both provide remarkably higher net throughput than the system with fixed clustering.**

***Index Terms*—User-centric, training resource allocation, base station cooperation transmission**

I. Introduction

With the trend of network densification [1], user-centric is becoming one of the design goals for fifth generation (5G) cellular systems. While there exist various meanings for "user-centric" [2, 3], one implication is that no matter where a user is located, its quality of service (QoS) requirement can be satisfied. To achieve such an ambitious goal, many techniques can be employed, e.g., a specific beam can be formed for a user in massive multi-input multi-output (MIMO) systems. When the base stations (BSs) in a dense network can share information via backhaul, another natural way is allowing each user to select several preferred BSs for transmission in a coordinated manner according to its QoS requirement and channel condition. Such BS cooperation networks are referred to as *user-centric networks* in this work.

Inter-cell interference (ICI) is a limiting factor to improve QoS in multicell networks, especially for the users located in cell-edge. BS cooperation transmission, either network MIMO or coordinated beamforming (CB), is effective to reduce ICI. Considering the backhaul limitation and training overhead in realistic networks, it is highly desirable for a cooperative cluster only consisting of few BSs [4]. In traditional cooperative networks, the clusters are formed from the perspective of network to improve spectral efficiency, which is non-overlapped [5], thereby the QoS of the cluster-edge users will be severely degraded by the out-of-cluster interference (OCI) [6]. In user-centric networks, the clusters are formed from the perspective of each user to ensure QoS. This inevitably leads to overlapped clusters [7, 8], as illustrated in Fig. 1.

For such user-centric networks, many challenges arise in the design for clustering, precoding, and training for channel estimation, due to the conflicting requirements from multiple users with different views. To resolve the conflict, a center unit (CU) is necessary. In reality, the CU can be deployed under the architecture of cloud radio access networks (C-RAN), which is recently proposed as a pivotal technology for future 5G [3]. CRAN enables centralized management and processing for large scale networks, where the baseband units are migrated from the BSs to a central entity. Nonetheless, computational complexity and signaling overhead become critical issues when the network grows to large. Specifically, to compute joint precoding for multiple BSs where each BS may belong to different clusters simultaneously, the instantaneous channel information from all BSs to all users within the cooperative clusters in the whole network should be gathered at the CU, which yields large training overhead for time-division duplex (TDD) systems or large training and feedback overhead for frequency-division duplex (FDD) systems if not judiciously designed.

The performance of BS cooperative transmission largely depends on the channel quality. To optimize the channel estimation that facilitates downlink precoding in cooperative networks, the training signals sent from the BSs in the same cluster or from the users selecting the same BS should be mutually orthogonal, which has been widely recognized and already reflected in 3GPP specification [9]. For non-overlapped clusters, this can be implemented by simply assigning a group of orthogonal training resources to each cluster and reuse the resources among clusters. For overlapped clusters, however, the orthogonal requirements for one cluster is coupled with another. To satisfy the requirements of all clusters simultaneously, a straightforward way is to assign each BS or each user an orthogonal training resource (e.g., scheme 1 in Fig. 1), whose overall training overhead increases linearly with the total number of BSs or users, which will counteract the cooperation gain and lead to low net throughput. By properly designing the reuse scheme of training resource among BSs or users, the overhead can be largely reduced (e.g., scheme 2 in Fig. 1).

Training signal design problem has been extensively inves-

Z. Chen and C. Yang are with the School of Electronics and Information Engineering, Beihang University, Beijing 100191, China, e-mail: {zchen,cyyang}@buaa.edu.cn, and X. Hou is with Department of Wireless Communications, China Mobile Research Institute, Beijing 100053, China e-mail:houxueying@chinamobile.com

This work was supported by the National 863 Program 2014AA01A703.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

tigated in literature, which focus on the design of training sequences, training duration, pilot placement and power allocation, mainly for single cell single user systems, e.g., [10] and references therein. If the OCI, ICI and multi-user interference in channel estimation can be avoided by assigning orthogonal training resources, all these well-explored techniques can be applied for user-centric networks. Nonetheless, how to allocate orthogonal training resource without causing heavy training overhead for a multi-cell network with overlapped cooperative clusters still remains unknown, which is important to improve the spectral efficiency of user-centric networks.

The goal of this paper is to reduce training overhead for downlink user-centric networks. To this end, we assign training resources to minimize overall training overhead of the network under the orthogonality constraint, which can realized by frequency division, time division or code division multiplexing in practice. We first describe the problems of downlink and uplink training resource allocation respectively for FDD and TDD systems in a general framework, and then focus on the downlink training scenario for easy exposition. We obtain the optimal solution with a graph-theoretic approach. To be scalable in terms of computational complexity, we propose a low complexity algorithm. Simulation results show that the algorithm performs close to the optimal solution, and both support higher net throughput than the schemes with fixed clustering.

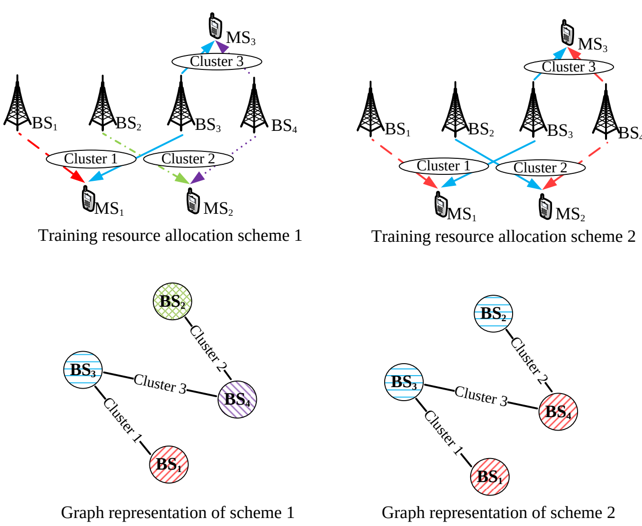

Fig. 1. An example of training resource allocation and its graph representation in a downlink user-centric network, where $B = 4$ and $K = 3$, $\mathcal{C}_1 = \{1,3\}$, $\mathcal{C}_2 = \{2,4\}$, $\mathcal{C}_3 = \{3,4\}$. Different types of lines represent different time slots used for training. In scheme 1, a total of four time slots are used, while in scheme 2, only two time slots are used.

II. System model

Consider a downlink multicell MIMO network, where K mobile stations (MSs) are located in B cells. The index sets of the BSs and MSs are denoted as $\mathcal{B} = \{1, 2, \ldots, B\}$ and $\mathcal{K} = \{1, 2, \ldots, K\}$, respectively. Each BS is equipped with N_t antennas and each MS is with N_r antennas.

For the user-centric networks under investigation, each user can select several BSs to form a cooperation cluster, say MS_k selects cluster $\mathcal{C}_k \subseteq \mathcal{B}$, in which the BSs can transmit to the MS either jointly with network MIMO or individually with CB. Since multiple MSs have different QoS requirements and channel conditions, their preferences will differ. This inevitably leads to overlapped clusters, i.e., $\mathcal{C}_k \cap \mathcal{C}_j \neq \emptyset, k, j \in \mathcal{K}, k \neq j$. In other words, each BS may belong to different clusters simultaneously, see Fig. 1 for example.

When the clusters are changed, the updated clustering results need to be sent to the CU for coordination. In order not to cause heavy signaling overhead, it is highly desirable to form the clusters with average channel gains in practice. Specifically, each MS can first measure and then report its average channel gains from several neighboring BSs to the CU, and the CU decides the cluster for each MS according to the available system resource and the preference of each MS [11]. In this way, the frequency of sending information to the CU depends on the change of user location, which is on the scale of second.

To facilitate downlink transmission, the channel information should be obtained at the BSs either from downlink training and uplink feedback in FDD systems or by uplink training in TDD systems. To capture the essence of the orthogonal training resource allocation problem, we consider single carrier systems with time-division orthogonality, where a basic training resource is a time slot. Nonetheless, the proposed method can be easily extended to multicarrier systems without fundamental difference, where the orthogonality can be ensured either via time- or frequency- or code-division or all of them simultaneously. In each downlink or uplink frame with overall M time slots, $M_{\text{tr}}^{\text{DL}}$ or $M_{\text{tr}}^{\text{UL}}$ time slots among them are used for downlink or uplink training, and the remaining time slots are used for downlink or uplink data transmission. The number of time slots used for training, i.e., $M_{\text{tr}}^{\text{DL}}$ or $M_{\text{tr}}^{\text{UL}}$, is referred to as the *training overhead.*

For downlink training in FDD systems, the channels from the BSs in cluster $\mathcal{C}_k$ to MS_k are estimated at MS_k. The training signals received at MS_k can be expressed as

$$\mathbf{Y}_k^{\text{DL}} = \sum\nolimits_{i \in \mathcal{C}_k} \mathbf{H}_{ki}\mathbf{X}_i^{\text{DL}} + \sum\nolimits_{i \notin \mathcal{C}_k} \mathbf{H}_{ki}\mathbf{X}_i^{\text{DL}} + \mathbf{N}_k^{\text{DL}} \quad (1)$$

where $\mathbf{H}_{ki} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix from BS_i to MS_k, $\mathbf{X}_i^{\text{DL}} \in \mathbb{C}^{N_t \times M_{\text{tr}}^{\text{DL}}}$ is the training matrix sent from BS_i, and $\mathbf{N}_k^{\text{DL}}$ is the noise matrix. The first term is the signal from the BSs in the cluster $\mathcal{C}_k$ selected by MS_k, and the second term is the interference from the BSs outside $\mathcal{C}_k$.

For uplink training in TDD systems, the channels from the users who select BS_b in their clusters are estimated at BS_b. The training signal received at BS_b can be expressed as

$$\mathbf{Y}_b^{\text{UL}} = \sum\nolimits_{i, b \in \mathcal{C}_i} \mathbf{H}_{ib}^{\text{T}}\mathbf{X}_i^{\text{UL}} + \sum\nolimits_{i, b \notin \mathcal{C}_i} \mathbf{H}_{ib}^{\text{T}}\mathbf{X}_i^{\text{UL}} + \mathbf{N}_b^{\text{UL}} \quad (2)$$

where $\mathbf{H}_{ib} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix from MS_i to BS_b, $\mathbf{X}_i^{\text{UL}} \in \mathbb{C}^{N_r \times M_{\text{tr}}^{\text{UL}}}$ is the training matrix sent from MS_i, $\mathbf{N}_b^{\text{UL}}$ is the noise matrix, and $(\cdot)^{\text{T}}$ denotes the transpose. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

first term is the signal from the MSs who select BS_b, and the second term is the interference from other MSs.

Comparing (1) and (2), we observe that the received signals in downlink and uplink are similar. For easy understanding and conciseness, we only consider downlink training in the rest of the paper, but the results can be easily extended to uplink training. In this way, we can drop the subscripts "DL" and "UL".

For OCI-free MIMO systems, the training signals among multiple co-located antennas at each BS should be mutually orthogonal to minimize the mean square error of channel estimation [12]. For network MIMO or CB, similarly, the training signals from different BSs in a cluster should also be orthogonal. The orthogonal training constraints can be expressed as: (1) $\mathbf{X}_i\mathbf{X}_i^{\text{H}} = p_t\mathbf{I}$, and (2) $\mathbf{X}_i\mathbf{X}_j^{\text{H}} = \mathbf{0}$ for $i, j \in \mathcal{C}_k, i \neq j$, where p_t is the transmit power at each BS, $\mathbf{I}$ is an identity matrix, $\mathbf{0}$ is a zero matrix and $(\cdot)^{\text{H}}$ is the conjugate transpose. For the considered user-centric network, the OCI should be controlled to improve channel estimation.

III. Optimal Orthogonal Resource Allocation

Due to the overlapped clusters, the resource allocation assignment for multiple BSs can not be decoupled. To minimize the overall training overhead subject to the orthogonal constraints, the training resource allocation needs to be coordinated from the network level at the CU. In this section, we formulate the optimal training resource allocation problem and find an optimal solution.

A. Problem Formulation

Assume that the cluster of each user has been formed based on its average channel gains. The clustering results can be represented by a BS-MS association matrix $\mathbf{A}$ of size $K \times B$, whose kb^{th} element $a_{kb} \triangleq 1$ if $b \in \mathcal{C}_k$ and $a_{kb} \triangleq 0$ if $b \notin \mathcal{C}_k$, where "1" or "0" indicates whether BS_b is selected by MS_k or not.

Denote $\mathbf{c}_b \triangleq [a_{1b}, \dots, a_{Kb}]^{\text{T}}$ as the b^{th} column vector of $\mathbf{A}$. Then, the value of $\mathbf{c}_i^{\text{T}}\mathbf{c}_j = \sum_{k\in\mathcal{K}} a_{ki}a_{kj}, i, j \in \mathcal{B}, i \neq j$ indicates whether BS_i and BS_j cooperate with each other. Since both a_{ki} and a_{kj} are binary variables, $\mathbf{c}_i^{\text{T}}\mathbf{c}_j = 0$ if and only if $a_{ki}a_{kj} = 0$ for all $k \in \mathcal{K}$, which reflects the fact that no MSs select both BS_i and BS_j (i.e., BS_i does not cooperate with BS_j). If $\mathbf{c}_i^{\text{T}}\mathbf{c}_j > 0$, there exists at least one user $k \in \mathcal{K}$ that makes $a_{ki}a_{kj} = 1$, which indicates that BS_i and BS_j are in the cooperative cluster of MS_k. Based on the value of $\mathbf{c}_i^{\text{T}}\mathbf{c}_j$, we define a BS-BS association matrix $\mathbf{F}$ of size $B \times B$, whose $ij^{\text{th}}(i \neq j)$ element $f_{ij} \triangleq 1$ if $\mathbf{c}_i^{\text{T}}\mathbf{c}_j > 0$ and $f_{ij} \triangleq 0$ if $\mathbf{c}_i^{\text{T}}\mathbf{c}_j = 0$, where "1" or "0" indicates whether BS_i and BS_j cooperate with each other. For $i = j$, we define $f_{ij} = 0$. An example of the matrices $\mathbf{A}$ and $\mathbf{F}$ is shown in Fig. 2.

Denote the index set of the time slots in each downlink frame as $\mathcal{M} \triangleq \{1, 2, \dots, M\}$. Suppose that the the training signal from each antenna is transmitted only in one time slot. After the CU allocating a time slot to each antenna, the results can be expressed by an allocation matrix $\mathbf{S}$ of size $B \times N_t$, whose bp^{th} element s_{bp} indicates the index of the time slot allocated to the p^{th} antenna at the b^{th} BS. Then, the number of different elements in the matrix $\mathbf{S}$ is the training overhead, M_{tr}.

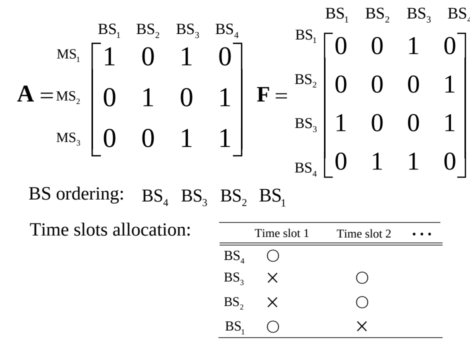

Fig. 2. The BS-MS and BS-BS association matrices for the network in Fig. 1, and the corresponding training resource allocation for each BS and the BS order selection obtained by *Algorithm 1*. We set $G = 2$. Then, $M_{\text{LB}} = 2$ is used in the optimal solution. Denote $\mathcal{B}_{16} = \{1, 2, 3, 4\}$. The recursive procedure to obtain the optimal solution begins with M_{16}. A possible division is $\mathcal{B}_{16} = \{1, 4\} \cup \{2, 3\}$ where $\{1, 4\}$ satisfies (7) and (8). This recursion continues as $\{2, 3\} = \{2, 3\} \cup \emptyset$ where $\{2, 3\}$ satisfies (7) and (8). Then the branch ends and $\mathcal{B}_{16}$ is divided into two disjoint subsets $\{1, 4\}$, $\{2, 3\}$ with each of them satisfying (7) and (8). Hence, $M' = 2$ and $M' = M_{\text{LB}}$. The recursion is terminated, and the minimal number of time slots for training is $M_{16} = 2$. The mark "×" indicates that allocating the corresponding time slot to the corresponding BS does not satisfy the constraint in (10) and (11), and the mark "◯" denotes that the time slot is allocated to the BS for training. By using *Algorithm 1*, the BS order is arranged as BS_4, BS_3, BS_2, BS_1. Recalling the definition of $\mathbf{s}$, the allocation vector is $\mathbf{s} = [1\,2\,2\,1]^{\text{T}}$, and the training overhead is $M_{\text{tr}} = 2$.

To satisfy the inter-antenna orthogonality, i.e., $\mathbf{X}_i\mathbf{X}_i^{\text{H}} = p_t\mathbf{I}$, the time slots assigned to multiple antennas at each BS, say BS_b, should be different, i.e.,

$$s_{bp} \neq s_{bq}, \;\; \text{if} \;\; p \neq q, p, q \in \{1, \cdots, N_t\} \tag{3}$$

To satisfy the inter-BS orthogonality, i.e., $\mathbf{X}_i\mathbf{X}_j^{\text{H}} = \mathbf{0}$ for $i, j \in \mathcal{C}_k, i \neq j$, the time slots assigned to BS_i and BS_j within the same cluster should differ, i.e.,

$$s_{ip} \neq s_{jq}, \;\; \text{if} \;\; f_{ij} = 1 \tag{4}$$

To reduce the training overhead, we allow each time slot to be reused for the BSs in different clusters. If these BSs interfere with each other, then OCI exists in training. Since the OCI is caused by transmitting training signals in the same time slot and the number of strong interference is finite with the user-centric clustering, a simple but efficient way to reduce OCI is to control the reuse times of each time slot, i.e.,

$$\max_m g_m \leq G \tag{5}$$

where g_m represents the times of using of the m^{th} time slot, $m \in \mathcal{M}$ (for example, $g_m = 0$ indicates that the m^{th} time slot is never allocated for training), G is a parameter that can be predetermined according to the practical interference environment, such as the density of the BSs.

To satisfy the requirements (3)-(5) in the user-centric network, an immediate approach inheriting the training design

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

criterion in prevalent systems is to allocate $N_t \cdot B$ different time slots from the index set $\mathcal{M}$ to the B BSs. With such an approach called as *straightforward training resource allocation method* in the rest of this paper, the training overhead is $M_{\rm tr} = N_t \cdot B$. Yet by properly reusing the time slots among the BSs, the overhead can be much smaller. This can be achieved by assigning training resource to minimize the training overhead under constraints (3), (4), and (5). For notational simplicity but without loss of generality, we first consider single-antenna case, and then extend to multi-antenna scenarios. When $N_t = 1$, the inter-antenna orthogonality need not to be considered, and only one time slot is required for each BS. Then, the $B \times N_t$ allocation matrix $\mathbf{S}$ is degenerated to a $B \times 1$ allocation vector $\mathbf{s}$, and the optimization problem is

$$\min_{\mathbf{s}} \quad M_{\rm tr} \quad s.t. \quad (4), (5) \tag{6}$$

B. Optimal Training Resource Allocation

Problem (6) is a kind of resource assignment problem, which has been widely modeled as a graph coloring problem (e.g. see [13–15] and references therein). By using each vertex of the graph to represent each BS, each edge in the graph to represent a conflict between two cooperated BSs, and each color to represent each time slot for training, problem (6) can be equivalently formulated as the following coloring problem.

P: *Coloring the B vertices in the graph with minimal number of colors under two constraints:*

(a). *Any two connected vertices should have different colors* (corresponding to constraint (4)),

(b). *Each color should not be used more than G times* (corresponding to constraint (5)).

Although graph coloring has been extensively studied, the problem we investigated differs from previous problems either in the objective or in the constraints. Our problem aims at minimizing the training overhead, whereas the resource assignment problems in [13–15] aimed at maximizing the capacity or the resource utilization efficiency. Our problem has an additional constraint (b) imposed to control the number of vertices sharing the same color, whereas the classical coloring problem in graph theory [16] aimed to minimize the number of colors only has constraint (a). Consequently, the algorithm to find the optimal solution proposed in [16] cannot be applied. To be specific, the algorithm in [16] sequentially partitions the original vertex set and its subsets into *maximal independent sets*, and colors each of them with one color to find the least number of colors. In problem **P**, these vertices may no longer be partitioned into *maximal independent sets*, hence a different way to partition the vertices is required.

After reformulating problem (6) as problem **P**, the index sets of the B BSs in the network $\mathcal{B}$ becomes the index set of the B vertices in the graph, and $f_{ij} = 1$ indicates that there is an edge between vertices $i, j \in \mathcal{B}$. Denote all 2^B subsets of $\mathcal{B}$ as $\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_{2^B}$, where $\mathcal{B}_1 = \emptyset$, $\mathcal{B}_{2^B} = \mathcal{B}$, and $|\mathcal{B}_i| \leq |\mathcal{B}_j|$ if $i < j$, i.e., the cardinalities of the subsets are in a non-descending order. Instead of coloring the vertices in $\mathcal{B}$ directly, we color the vertices in subset $\mathcal{B}_i$. Denote M_i as the minimal number of colors that are required to color the vertices in $\mathcal{B}_i$ under constraints (a) and (b), and set $M_1 = 0$. The following proposition indicates that the optimal solution of problem (6) or problem **P**, M_i, can be obtained in a recursive manner.

Prop. 1: Divide $\mathcal{B}_i$ ($i > 1$) into two disjoint subsets $\mathcal{B}_i = \mathcal{B}_{i_1} \cup \mathcal{B}_{i_2}$, where the nonempty $\mathcal{B}_{i_1}$ satisfies

$$f_{jk} = 0, \quad j, k \in \mathcal{B}_{i_1} \tag{7}$$

$$|\mathcal{B}_{i_1}| \leq G, \tag{8}$$

then the minimum number of colors required by arbitrary vertex set $\mathcal{B}_i$ is

$$M_i = \min_{i_2} M_{i_2} + 1 \tag{9}$$

Proof: Assume that the minimum of M_i is obtained when $i_2 = i_2'$, and then $M_i = M_{i_2'} + 1$.

First, we show that the vertices in arbitrary vertex set $\mathcal{B}_i$ can be colored with $M_{i_2'} + 1$ colors. Note that (7) means that any two vertices in the 1^{st} subset $\mathcal{B}_{i_1}$ are not connected, and (8) means that the number of vertices in $\mathcal{B}_{i_1}$ is no more than G. If we color all vertices in $\mathcal{B}_{i_1}$ with one color, both constraints (a) and (b) will be satisfied. Therefore, $M_{i_1} = 1$, and in total $M_{i_2'} + 1$ colors are able to color all the vertices in $\mathcal{B}_i$.

Second, we prove that $M_{i_2'} + 1$ is the minimal number of colors by contradiction. Suppose that the vertices in $\mathcal{B}_i$ can be colored with $\tilde{M} < M_{i_2'} + 1$ colors. Based on the coloring result, we divide $\mathcal{B}_i$ into two disjoint subsets as $\mathcal{B}_i = \mathcal{B}_{i_1''} \cup \mathcal{B}_{i_2''}$, where $\mathcal{B}_{i_1''}$ consists of the vertices with one color, and $\mathcal{B}_{i_2''}$ consists of the remaining vertices with other $\tilde{M} - 1$ colors. Since the vertices in the 1^{st} subset $\mathcal{B}_{i_1''}$ share one color, the vertices in $\mathcal{B}_{i_1''}$ are not connected, and $|\mathcal{B}_{i_1''}| \leq G$, which indicates that $\mathcal{B}_{i_1''}$ satisfies (7) and (8). Since the vertices in $\mathcal{B}_{i_2''}$ can be colored with $\tilde{M} - 1$ colors, $M_{i_2''} \leq \tilde{M} - 1$. Then $\min_{i_2} M_{i_2} \leq M_{i_2''} \leq \tilde{M} - 1 < M_{i_2'}$, which contradicts the assumption that the minimum of M_{i_2} is obtained at $i_2 = i_2'$ in the beginning. ■

This proposition suggests that we can find the optimal solution of problem **P** (and hence problem (6)) recursively by starting from partitioning the set $\mathcal{B}$. To obtain (9), all possible divisions should be enumerated. The prohibitive recursion branches can be reduced in the considered problem by a terminating condition. Note that each branch ends when $\mathcal{B}_{i_2} = \emptyset$, i.e., $M_{i_2} = 0$. Suppose that when a branch ends the set $\mathcal{B}$ has been divided into M' disjoint subsets with each of them satisfying (7) and (8). From (5) and noting that $\sum_m g_m = B$ when $N_t = 1$, we know that $M_{\rm tr} \geq \lceil B/G \rceil \triangleq M_{\rm LB}$, where $\lceil(\cdot)\rceil$ is the smallest integer no less than $(\cdot)$. Then, if $M' = M_{\rm LB}$, we can terminate the recursion since the lower bound is achieved. Otherwise, the recursive procedure continues. To help understand the procedure, an example for is provided in the legend of Fig. 2.

Complexity Analysis: The major complexity of the algorithm lies in enumerating all possible divisions, which is huge when G is large. As a rule of thumb estimation, the complexity

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

scales with the network size as $O(2^B)$ [17] (scales with the number of users as $O(2^K)$ in uplink training case.)

Extend to Multiple-antenna Case: When each BS has multiple antennas, each vertex in the graph corresponds to each antenna, and the edge corresponds to the orthogonality constraints between two antennas. Then, the recursive procedure can be applied to find the optimal training resource allocation.

IV. Low-complexity Training Resource Allocation Algorithm

In this section, we propose a low-complexity algorithm to find a suboptimal solution of problem (6) for the single-antenna case. The basic idea is to arrange all BSs in an order and allocate the time slots for training to each BS sequentially, where the BS order and resource allocation are respectively optimized. While the idea to reduce complexity by sequential allocation is also applied for deriving heuristic algorithms of other coloring problems [13, 15, 18], the existing algorithms are not applicable to our problem since problem **P** differs from previous problems either in the objective function or the constraints.

A. Training Resource Allocation to Each BS

We first optimize the training resource allocation for a given BS order. Suppose that all B BSs have been arranged in a queue as $\mathrm{BS}_{n_1}, \cdots, \mathrm{BS}_{n_B}$, where indices $\{n_1, \ldots, n_B\}$ are integers within the set $\mathcal{B}$. We start the procedure by allocating a time slot with index in $\mathcal{M}$, say the 1^{st} time slot, to BS_{n_1}. Denote s_{n_1} as the n_1^{th} element of the $B \times 1$ allocation vector $\mathbf{s}$, and then $s_{n_1} = 1$. Next we select a time slot from $\mathcal{M}$ for BS_{n_2}. This procedure continues for each $\mathrm{BS}_{n_3}, \ldots, \mathrm{BS}_{n_B}$ in a sequential manner. When we select the m^{th} time slot from $\mathcal{M}$ for $\mathrm{BS}_{n_i}, i \geq 2$, the inter-BS orthogonal constraint in (4) and the maximum reuse time limitation in (5) should be satisfied, i.e.,

$$m \neq s_j, \ \text{if} \ f_{jn_i} = 1, \tag{10}$$

$$g'_m < G \tag{11}$$

where $j \in \{n_1, \cdots, n_{i-1}\}$ is the index of the BS that arranged ahead of BS_{n_i}, s_j is the index of the time slot allocated to BS_j, and g'_m is the times of using the m^{th} time slot among previous $i-1$ allocations.

There may exist more than one time slot, say the one with index $m \in \mathcal{M}$, that satisfies (10) and (11). Among them we select the time slot, say with index m', that has been used before but with the minimum reuse times, i.e.,

$$m' = \arg \min_{m, g'_m \neq 0} g'_m \tag{12}$$

in order to make the reuse times of different time slots as uniform as possible, which helps reduce training overhead. After m' is obtained, we allocate the m'^{th} time slot to BS_{n_i}, which is represented by $s_{n_i} = m'$. Note that such a criterion to select color for each BS differs from the existing algorithms [18] due to the extra constraint. To help understand, again we use the example network in Fig. 2 to illustrate the procedure.

B. Order Selection for BSs

Next, we optimize the BS order to minimize the training overhead. Since closed-form expression of M_{tr} is hard to derive, we provide a simple upper bound of M_{tr} for the sequential resource allocation that depends on the BS order, and then select the BS-order that minimizes the upper bound. Note that for any given order, $\mathrm{BS}_{n_1}, \cdots, \mathrm{BS}_{n_B}$, there are $i-1$ BSs arranged ahead of BS_{n_i}. Among the $i-1$ BSs, the number of BSs that cooperate with BS_{n_i} is

$$F_i \triangleq \sum\nolimits_{j \in \{n_1, \cdots, n_{i-1}\}} f_{jn_i} \tag{13}$$

and the rest of $i-1-F_i$ BSs do not cooperate with BS_{n_i}.

Consider the worse case where the time slots allocated to the F_i cooperative BSs are totally different. To satisfy (10), which indicates that the time slot allocated to BS_{n_i} should be different from those allocated to its cooperative BSs, we should allocate a new time slot, which is different from the F_i time slots, to BS_{n_i}. Furthermore, considering that the new time slot may also be used by the $i-1-F_i$ non-cooperative BSs, to satisfy (11) at most $\lfloor (i-1-F_i)/G \rfloor + 1$ new time slots are required, where $\lfloor (\cdot) \rfloor$ is the largest integer not greater than $(\cdot)$. Therefore, at most $F_i + \lfloor (i-1-F_i)/G \rfloor + 1$ time slots are required to satisfy both (10) and (11). Then, an upper bound of M_{tr} can be obtained as

$$M_{\text{tr}} \leq \max_{i \in \mathcal{B}} (F_i + \lfloor (i-1-F_i)/G \rfloor + 1) = \max_{i \in \mathcal{B}} (\lfloor \frac{G-1}{G} F_i + \frac{i-1}{G} \rfloor + 1) \tag{14}$$

from which we can see that the upper bound can be reduced by reducing F_i. This can be achieved by arranging the order of the BSs in the queue. Based on (13), recalling that f_{jn_i} is a binary variable, we obtain $F_i \triangleq \sum_{j \in \{n_1, \cdots, n_{i-1}\}} f_{jn_i} < \sum_{j \in \{n_1, \cdots, n_{i-1}\}} 1 = i-1$ and $F_i \triangleq \sum_{j \in \{n_1, \cdots, n_{i-1}\}} f_{jn_i} < \sum_{j \in \mathcal{B}} f_{jn_i}$, then

$$F_i \leq \min\{C_{n_i}, i-1\} \tag{15}$$

where $C_{n_i} \triangleq \sum_{j \in \mathcal{B}} f_{jn_i}$ is the number of cooperative BSs of BS_{n_i}. (15) suggests that if we design the order $n_1, n_2, \cdots, n_B$ to make $C_{n_1} \geq C_{n_2} \geq \cdots \geq C_{n_B}$, then all the values of $F_1, F_2, \cdots, F_B$ will be small. The intuitive meaning of such an order selection is first allocating training resource to the BS who belongs to most clusters. The training resource allocation algorithm is summarized in *Algorithm 1*, which includes the BS order selection and the resource allocation for each BS.

Complexity Analysis: To compute $C_i, i \in \mathcal{B}$ and sort C_i in a descending order in Step 1 of the algorithm, at most $(B-1)B$ integer additions and $(B-1)B$ comparisons are required. Thus, the complexity to select the order for the BSs is $\mathcal{O}((B-1)B + (B-1)B) = \mathcal{O}(B^2)$. To select a time slot that satisfies constraints (10) and (11) for each BS, at most $B-1$ comparisons are required. Hence, for the overall B BSs in total $(B-1)B$ comparisons are required for the allocation. Thus the complexity scales with the network size as $\mathcal{O}(B^2 + (B-1)B) = \mathcal{O}(B^2)$ (scales with the number of users as $\mathcal{O}(K^2)$ in the uplink training case).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

Algorithm 1 Low complexity training resource allocation algorithm

Step 1: Sort the BSs in the order $\text{BS}_{n_1}, \ldots, \text{BS}_{n_B}$ with $C_{n_1} \geq C_{n_2} \geq \ldots \geq C_{n_B}$.
Step 2: Initialize the algorithm by allocating the 1^{st} time slot to BS_{n_1} and setting $i = 2$.
Step 3: Select the m^{th} time slot from $\mathcal{M}$ for BS_{n_i}, where m should satisfy (10) and (11). If there are more than one time slot satisfying (10) and (11), we choose the one obtained from (12).
Step 4: Update $i = i + 1$. Return to Step 3 until all BSs are allocated with time slots.

Extension to Multiple-antenna Case: For the BSs with N_t antennas, we should select N_t time slots for each BS, and each of the N_t time slots should satisfy the inter-BS orthogonal constraint and the maximum reuse times constraint. When there exist multiple time slots that satisfy both constrains in each selection, we finally choose the one that has been used before but with the minimum reuse times as in (12).

V. Simulation and Numerical Results

In this section, we evaluate the performance of the proposed training resource allocation algorithm. Since the resource allocation for uplink and downlink training is similar, we provide the results for uplink training assisting TDD downlink transmission to avoid channel feedback issues. Consider a network consisting of B macro cells, where B = 19 or 7. To remove boundary effect, we consider wrap around at the boundary of the B cells, then any cell can be regarded as being surrounded by the same tiers of cells. The cell radius is 250 m. The path-loss is modeled as $\text{PL}^{\text{dB}} = 35.3 + 37.6\log_{10}(d_{k,b})$ [9], where $d_{k,b}$ (in meter) is the distance between BS_b and MS_k. The transmit power of each BS is set as $p_t = 46$ dBm and the noise power is -95 dBm. Considering the difference of the transmit powers between uplink and downlink, we set the uplink receive signal-to-noise ratio (SNR) 5 dB lower than the downlink SNR. A simple clustering method based on the average channel gain differences of each user is employed. Specifically, denote α_{th} as a threshold for the average channel gain difference used to control cluster size, and $\alpha_k^{\max}$ as the strongest average channel gain of MS_k. If from a BS to MS_k the average gain exceeds $\alpha_k^{\max} - \alpha_{\text{th}}$, the BS will belong to the cluster of MS_k. The results are obtained by averaging over 500 trials, where in each trial the users each with a single antenna are randomly placed. Unless otherwise specified, these parameters are used in the sequel.

In Figs. 3 and 4, we compare the average number of allocated time slots for uplink training achieved by the low-complexity algorithm (with legend "User-centric seri.") and the optimal solution (with legend "User-centric opt."), respectively versus α_{th} and G with different values of K and B. We can see that the overall training overhead of the low complexity algorithm is almost the same as the optimal solution. As the value of α_{th} grows, the cluster size increases and hence the training overhead grows correspondingly. As the value of G grows, each time slot is reused more often and hence the overhead reduces until the constraint in (5) is not longer active. The training overhead almost linearly increases with α_{th} and the increasing rate is higher for larger value of K, but only increases a little with B if G is properly selected.

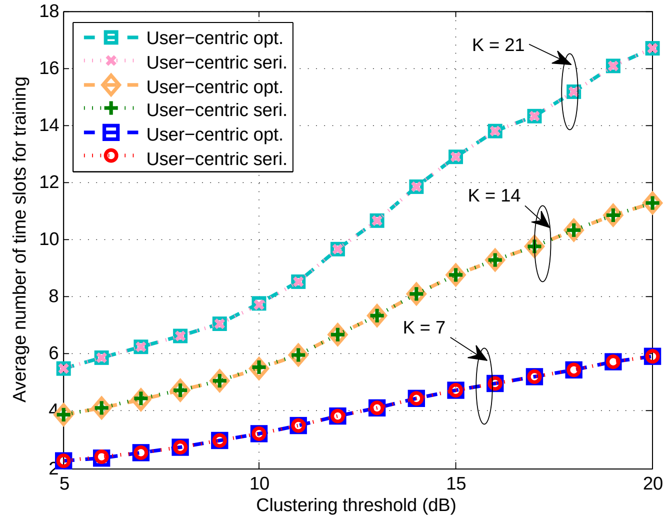

Fig. 3. Overall training overhead of the optimal solution and low complexity algorithm of training resource allocation versus α_{th}, $N_t = 1$, $N_r = 1$, $B = 7$, $G = 7$. When "User centric strait." scheme is used, the training overheard are 21, 14 and 7 respectively for the cases of K = 21, 14 and 7, which are independent from α_{th}. Note that these result are for uplink training, the overhead linearly increases with K, but only slightly increases with B, which is not shown due to the space limitation.

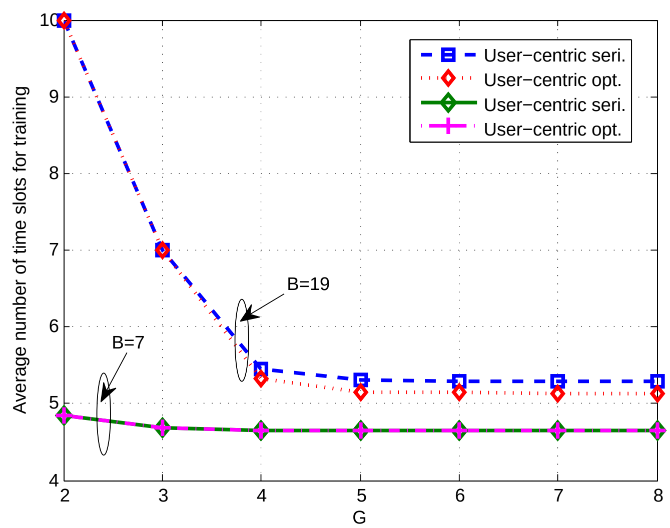

Fig. 4. Overall training overhead of the optimal solution and low-complexity algorithm of training resource allocation versus G, $N_t = 1$, $N_r = 1$, $K = B$, $\alpha_{th} = 15$ dB.

In Fig. 5, we show the net throughput of user-centric downlink transmission networks, defined as $R_{\text{net}} = (1 - \gamma M_{tr}) \sum_{k \in \mathcal{K}} \log_2(1 + \text{SINR}_k)$, where $\gamma = 1/M$ is the ratio of the resources occupied by one time slot to that used by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2015.2420114, IEEE Transactions on Vehicular Technology

the whole network, and SINR_k is the receive SINR of MS_k. We set $\gamma \in [0, 1\%]$ according to the configuration of sounding reference signal in 3GPP [19]. Considering that there is no precoding available for network MIMO with overlapped clusters, we employ CB to remove ICI. Except for the proposed solutions and the *straightforward training resource allocation method* with legend "User-centric strai.", we also show the performance of CB with fixed clustering (with legend "Fixed clustering") and non-coordinated systems (with legend "No-coord."). In the fixed clustering scheme, the training resources are orthogonal within each cluster, and are reused in adjacent clusters. The training resource allocation in non-coordinated networks is set the same as in fixed clustering networks so that the uplink channel estimation is under the same level of interference. For a fair comparison, we set $G = 7$ for the case $B = 19$. We employ minimum mean square error (MMSE) criterion to estimate the uplink channels, where OCI is treated as noise. We use zero-forcing beamforming for both coordinated and non-coordinated downlink transmission. For CB, the beamforming vector of MS_k at BS_b is $\mathbf{w}_k = \frac{(\mathbf{I}-(\hat{\mathbf{H}}^{\text{UL}}_{b,k})^H(\hat{\mathbf{H}}^{\text{UL}}_{b,k}(\hat{\mathbf{H}}^{\text{UL}}_{b,k})^H)^{-1}\hat{\mathbf{H}}^{\text{UL}}_{b,k})\hat{\mathbf{H}}^T_{kb}}{\left\|(\mathbf{I}-(\hat{\mathbf{H}}^{\text{UL}}_{b,k})^H(\hat{\mathbf{H}}^{\text{UL}}_{b,k}(\hat{\mathbf{H}}^{\text{UL}}_{b,k})^H)^{-1}\hat{\mathbf{H}}^{\text{UL}}_{b,k})\hat{\mathbf{H}}^T_{kb}\right\|}$ where $\hat{\mathbf{H}}_{kb} \in \mathbb{C}^{N_r \times N_t}$ ($N_r = 1$) is the estimated channel between $\text{BS}_b \in \mathcal{C}_k$ and MS_k, and $\hat{\mathbf{H}}^{\text{UL}}_{b,k}$ is the estimated channel matrix from BS_b to all the users selecting BS_b except MS_k. Since with CB each BS needs to equip enough antennas to avoid strong interference within each cluster, we set $N_t = 8$. The results show that "User-centric seri." outperforms "Fixed clustering" and "No-coord" due to the reduced OCI both in uplink training and downlink transmission, and performs much better than "User-centric strai." as the ratio γ increases. When the cells are denser with smaller radius, the proposed solutions are more superior.

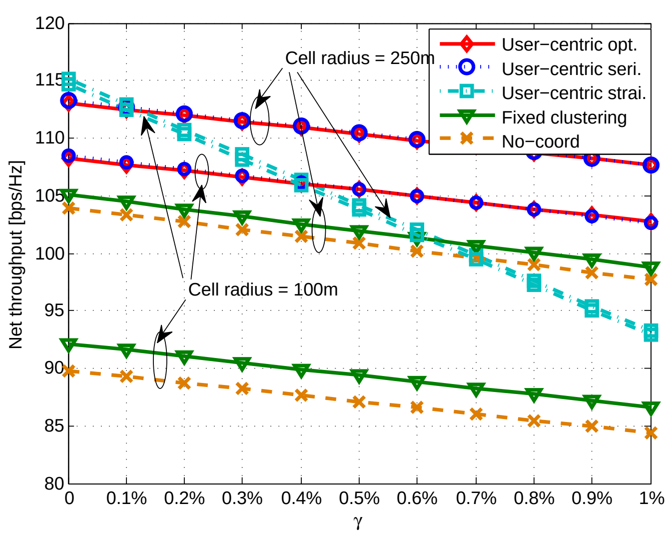

Fig. 5. Net throughput of the network versus the ratio γ, $N_t = 8, N_r = 1, B = 19, K = 19, G = 7$, the threshold for user-centric clustering is set as $\alpha_{th} = 15$ dB, the maximum size of each cluster is controlled as four that is a typical value in practical systems. In the fixed clustering scheme, five clusters of size three and two clusters of size two are formed.

In Fig. 6, we show how the net throughput scales with the network size. It is shown that the performance of all schemes increase linearly with B, except for the straightforward method due to large training overhead. The results of the scalability with K are similar, which are not provided due to the space limitation.

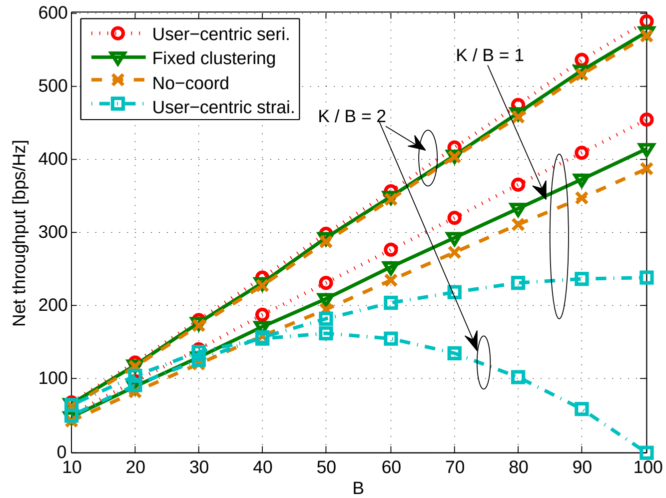

Fig. 6. Net throughput of the network versus B, $N_t = 8$, $N_r = 1$, $G = K/3$, $\gamma = 0.5\%$, $\alpha_{th} = 15$ dB, the maximum size of each cluster is controlled as four. To reflect a more realistic network, we randomly place the BSs in an area with average inter-BS distance of 500 m. In the fixed clustering scheme, the neighboring three BSs are formed as a cluster.

VI. Conclusions

In this paper, we studied training resource allocation problem for BS cooperation networks with overlapped clusters. We found the optimal solution to minimize training overhead under the inter-BS and inter-antenna orthogonal constraints and a constraint to control the OCI, and provide a low complexity algorithm for practice use. Simulation results show that the fast algorithm performs closely to the optimal solution, and both are superior to existing schemes with fixed clusters in terms of net throughput. The proposed algorithm is applicable to both TDD and FDD systems, and are easily extended to wideband systems where the orthogonality can be ensured either via time- or frequency- or code-division or all of them simultaneously.

References

[1] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[2] F. Boccardi, R. W. H. Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: a 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.

[4] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. J. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2983–2989, Jun. 2009.

[5] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in *Proc. IEEE ICC, 2008.*

[6] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inform. Theory*, vol. 59, no. 9, pp. 5213–5226, Sept. 2013.

[7] J. Gong, S. Zhou, Z. Niu, L. Geng, and M. Zheng, "Joint scheduling and dynamic clustering in downlink cellular networks," in *Proc. IEEE GLOBECOM, 2011.*

[8] A. Giovanidis, J. Krolikowski, and S. Brueck, "A 0–1 program to form minimum cost clusters in the downlink of cooperating base stations," in *Proc. IEEE WCNC, 2012.*

[9] 3GPP Long Term Evolution, "Further advancements for E-UTRA physical layer aspects," *TSG RAN TR 36.814 v9.0.0*, 2010.

[10] B. N. Bharath and C. R. Murthy, "Channel training signal design for reciprocal multiple antenna systems with beamforming," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 140–151, Jan. 2013.

[11] D. Liu, Q. Zhang, S. Han, C. Yang, G.Wang, and M. Lei, "Semi-dynamic cooperative cluster selection for downlink coordinated beamforming systems," in *Proc. IEEE WCNC 2014*.

[12] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[13] R. Y. Chang, Z. Tao, J. Zhang, and C.-C. J. Kuo, "Multicell OFDMA downlink resource allocation using a graphic framework," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3494–3507, Sept. 2009.

[14] Y.-S. Liang, W.-H. Chung, G.-K. Ni, I.-Y. Chen, H. Zhang, and S.-Y. Kuo, "Resource allocation with interference avoidance in OFDMA femtocell networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2243–2255, June 2012.

[15] K. Khalil, G. Farhadi, and A. Ito, "Iterative fair channel assignment for wireless networks," *IEEE Wireless Commu. Lett.*, vol. 3, no. 2, pp. 145–148, Apr. 2014.

[16] N. Christofides, "An algorithm for the chromatic number of a graph," *Computer Journal*, vol. 14, pp. 38–39, Feb. 1971.

[17] E. L. Lawler, "A note on the complexity of the chromatic number problem," *Inf. Process. Lett.*, vol. 5, no. 3, pp. 66–67, 1976.

[18] D. Brlaz, "New methods to color the vertices of a graph," *Commun. of the ACM*, vol. 22, no. 4, pp. 251–256, April 1979.

[19] 3GPP Long Term Evolution (LTE), "Physical channels and modulation," *TSG RAN TS 36.211 v10.5.0*, 2012.