Research on Radio Resource Management in the Next Generation of Cellular Network

September 2016

Jia Yu

Contents

Chapter 1 Introductory chapter

1.1 Research Backgrounds

As long with the development of mobile network techniques and the popularization of smart terminals, the demand for mobile communication traffic has been explosively increased. By the end of 2013, the average amount of digital traffic of global mobile communication system had been 1.5 EB per month, while the number at the same time of 2012 is 820 PB per month. At the end of 2014, this number had reached up to 2.5 EB per month [1]. As predicted by Cisco, digital traffic of global mobile communication system will grow up to 24.3 EB per month at the end of 2019. In order to meet the continuously increasing demand of digital traffic, higher requirements are proposed for the next generation of mobile communications systems. In March of 2008, the International Telecommunications Union-Radio communications sector (ITU-R) published a specification that determined the technical requirements for the Four Generation (4G) of mobile communication system [2]. As required, Instructions in the requirements, 4G system should be able to provide 100 Mbps mobile data transmission services for high-speed mobile users (such as trains, cars on the user), and provide 1 Gbps transmission services for the low-speed mobile users. As one of the 4G standard, Term Evolution-Advanced Long (LTE-A) system has been realized the peak data rate of 1 Gbps for downlink transmissions and 500 Mbps for uplink transmissions. Furthermore, higher technical requirements are proposed for the fifth Generation (5G) of mobile communication systems. European 5G summit in Munich held in February 2014 defined the targets of 5G, including 10 Gbps peak data rate, hundreds of Megabits per second for low-speed mobile users, more than ten Megabits per second for high-speed mobile users, 1000 times of system capacity and 10 to 100 times of device connections than the current networks.

With the increasing demand of mobile communication, the scarcity of spectrum resources is gradually exposed. Expensing new spectrum resources and improving the spectrum efficiency are two most promising methods for alleviating the spectrum-scarcity problem. To expanse spectrum resource for mobile communication systems, ITU-R proposes a detailed program about available frequency bands and bandwidths for different application scenarios, based on which each country further

programs more specifically according its own situation. According to the transmission characteristics of the wireless signal and the requirements of the mobile communication system, the frequency suitable for land mobile communication systems is in the band of High Frequency Ultra (UHF). However, in addition to the mobile communication system, the UHF is also shared by digital TV, Wi-Fi and Bluetooth and other wireless transmission techniques. At present, most of the UHF has been occupied; therefore, additional frequency that can be exploited for mobile communication system is not sufficient to meet the demands of data traffic in the near future.

In addition to expanding spectrum resources, improving the spectrum efficiency of mobile communication system is another effective way to alleviate the scarcity of spectrum resources. The spectral efficiency of the mobile communication system can be defined as the data rate per Hertz frequency (bit/s/Hz or bps/Hz). Although most of the existing mobile communication systems use advanced physical layer access technology, modulation coding technology to achieve spectral efficiency, but there are many other factors that affect the effective use of spectrum resources, such as intercell interference, inadequate resource allocation and so on.

The next generation mobile communication system must consider the serious energy consumption problem while pursuing improvement of spectral efficiency. According to the statistics in [3], the Information Communication Technology (ICT) industry led to 3% of the total global emissions of greenhouse gas. This means that ICT industry has been one of the major sources of environmental pollution around the world. In addition to causing serious environmental pollution, energy consumption also greatly increases the cost of operators. China Mobile c-ran project white paper [4] pointed out that energy consumption has become an important part of the operator costs. Among them, the energy consumption used for supporting sites is accounted for 24% of the total capital, and the annual electricity consumption of sites is accounted for up to 41%. Therefore, improving energy efficiency of the ICT industry is the requirement of not only social development, but also the development of mobile communications operators themselves. In order not to increase energy consumption, 5G mobile communication systems need to support more mobile data services at the current level of energy consumption. For example, if the 5G mobile communication system data rate increased by 100 times, then the energy consumption per bit should be 1/100 or smaller of existing system [5], that is, the system energy efficiency to be

increased by 100 times. However, there is a complex tradeoff between spectrum efficiency and energy efficiency in mobile communication systems. Under ideal conditions, energy efficiency and spectral efficiency is obviously in inverse proportion. In the practical system, if the circuit power is considered, the relationship between energy efficiency and spectral efficiency will be changed into a cup type curve [6]. In addition, the relationship between energy efficiency and spectral efficiency will become much more complex [7] when considering the imperfect channel state information (State Information Channel, CSI), signal synchronization costs and other similar problems. Improving spectrum efficiency by increase transmission power, as a traditional solution in mobile communication system, is no longer appreciated. How to greatly improve the spectrum efficiency while meeting the intense energy efficiency requirements of the system becomes a crucial problem for the next generation mobile communication system.

To meet the stringent requirements to the next generation mobile communication systems, several advanced techniques have been proposed. Carrier Aggregation (CA) is one of the most promising techniques to improve peak data rates. With CA, several contiguous or non-contiguous spectrum fragments are combined to serve a specific user, and the total bandwidth to the user can be up to 100 MHz in LTE-A systems if the user equipment is sufficiently capable. In addition, CA is also helpful to improve the spectrum efficiency of the network via making good use of non-contiguous carrier components. Another attractive method to improve spectrum efficiency of cellular networks is a new deployment widely known as Heterogeneous Networks (HetHets). A HetNet is a dense-deployed network where several low-cost low-power access points are added into a macro cell built by an eNodeB (eNB). Each access point is connected to the eNB by a dedicated backhaul link, and is capable to offload the heavy traffic load of the eNB. It is possible to obtain high-qualified transmissions between access points and users since propagation distances are shortened. Meanwhile, energy efficiency is improved in HetNets since power demanded for transmission is relatively lower than that of traditional cellular networks. However, the shortened distances between cells lead to exacerbated intercell interference, which effect the performance of networks in terms of spectrum efficiency. To deal with intercell interference Coordinated Multiple Points (CoMP) is introduced into LTE-A. CoMP is an efficacious solution to mitigate intercell interference by making Base Stations (BSs) or access points cooperate with each other.

Appropriate algorithms of Radio Resource Management (RRM) are essential for networks in practice. However, the techniques mentioned above significantly increase the complexity of RRM in a practice network. CA involves carrier components allocation which must be processed before further subcarrier and power allocation. HetNets and CoMP ask the RRM be conducted jointly among multiple BSs. As a result, RRM problem become NP-hard, and is impossible to achieve an optimal solution in polynomial time. In this case, it is necessary to pursue simplified algorithms of RRM for a network with these promising techniques, while maintain favorable network performances. The thesis focuses on the rationale behind CA, HetNet and CoMP, and dedicates to research on practical and efficacious algorithms of RRM in the network with these techniques.

1.2 State of the Art Related Work

Intercell interference is one of the major factors that damage the performance of networks, especially those are densely deployed. Therefore, RRM in a multicell scenario should consider reducing the effect caused by intercell interference. To this end, [8] suggest that the radio network controller (RNC) assign data traffic to each BS on the principle of minimizing intercell interference, and then each BS adaptively allocates subchannels to users. Similarly, in [9], the system selects a group of users with minimal intercell interference to communication. In addition, authors of [10] propose a Graph-theory-based scheduling algorithm to avoid intercell interference.

The application of the CA technique brings new challenges, especially in terms of RRM of LTE-A systems. The first concern is the computation complexity of resource allocation. There have been a few literatures on resource allocation with CA [11] [12] [13]. These works have a similar study framework that first decomposes the complex allocation problem into several subproblems, and then solves these sub-problems step by step. Although the decomposed framework may not get the optimal solution comparing with jointly allocation, it can significantly reduce computation complexity, which is of importance to a real-time system.

CoMP is beneficial to reduce intercell interference via allowing the coordination and cooperation between BSs. However, the complexity of RRM is significantly intensified. Authors of [14] propose a scheduling algorithm based on Signal-to-Leakage-plus-Noise Ratio (SLNR) for JP CoMP system. However, the limitation

caused by backhaul link is not considered. To fulfill CoMP transmissions, data packet and control information must be shared between BSs via backhaul links. Therefore, the capacity of backhaul links constrains the achievable data rates [15][16]. An adaptive precoding and power allocation algorithm is proposed in [17], where constrained backhaul capacity is considered. [18] proposes a system model that switch between CS/CB CoMP and JP CoMP: CS/CB CoMP is used when there are congestion on backhaul links, and JP CoMP is used in the other case.

Most references mentioned above are intended to improve spectrum efficiency of mobile systems. As the ICT has rapidly spread and upgraded, the absolute energy consumption caused by it becomes a big issue. Therefore, it is crucial to develop highly energy-efficient mobile systems. For this end, researches have widely studied on RRM with the purpose of improving energy efficiency. [19][20][21] build mathematical formulation of jointly allocating spectrum and power with the objective of minimizing total power allocation, and proposed RRM algorithm respectively. The authors of [22] and [23] consider the BS selection problem associating with power allocation. [24] limits the minimal data rate of each user, in order to avoid the unfavorable case where transmission quality dramatically decreases caused by energy saving. In addition, authors of [25] give a detailed review on energy efficiency RRM in multiuser cellular networks.

1.3 Main Contributions and Framework of the Thesis

This thesis focuses on RRM in mobile systems with promising technologies such as CA, HetNet and CoMP.

Chapter 2 introduces CA, HetNet and CoMP LTE-A system, which are considered in this thesis in the following chapters. The rationales behind these techniques are described to explain why they promising for significantly improve network performance.

Chapter 3 considers RRM in the wireless communication system with CA technique. We have decomposed the problem of resource allocation with CA into three subproblems, CC allocation, RB allocation, and power allocation. For each sub-problem, we propose low-complexity algorithms correspondingly.

Chapter 4 studies a constrained RRM problem being aimed at improving energy efficiency in a CoMP-based HetNet. To solve the problem, we first propose a CE-based

RB scheduling algorithm under the assumption of equal power allocation. Then a KKT-based algorithm for power allocation is proposed. The proposed algorithms are considered to be used in a centralized way at the first place. Since centralized strategy for RRM takes a long time delay in large-scale networks, we modified the proposed to adapt decentralized systems in order to shorten the time delay of RRM processing.

Chapter 2 Promising techniques in LTE-A: CA, HetNet and CoMP

Long Term Evolution-A is proposed by 3GPP on the base of LTE (3GPP Release 8/9), targeting at satisfying the requirements of IMT-Advanced. In addition to the previous techniques in LTE system, several promising techniques are introduced in LTE-A, such as CA, HetNet and (CoMP). This chapter explains the rationales behind these techniques in details.

2.1 Carrier Aggregation

CA is introduced by Rel-10 as a main feature of LTE-A systems for meeting the peak data rate requirements (1 Gbps and 500 Mbps for downlink and uplink, respectively) for 4G mobile communication systems [26]. CA combines spectrum component in continuous or non-continuous frequency bands to realize broadband transmission. Each individual frequency band used by CA is referred to as a Component Carrier (CC). The bandwidth of each CC could be 1.4, 3, 5, 10, 15 and 20 MHz, following the bandwidth configuration in LTE system. As specified in Rel-10, CA technically allows at most 5 CCs to be simultaneously used for a capable User Terminal (UE). This means that a bandwidth of up to 100 MHz can be achieved by aggregating 5 20MHz-CCs. In this way, peak data rates can be significantly improved.

To deal with different requirements and conditions of venders, the following three categories of CA techniques are defined by 3GPP [27]:

1) Intraband contiguous CA supports aggregation of adjacent CC in the same bandwidth. Intraband contiguous CA is easy to be deployed in the practical system. However, it is difficult to have several contiguous CCs in a signal frequency band due to the limitation of each band.
2) Intraband non-contiguous CA supports the aggregation of non-adjacent CCs in the same bandwidth. Frequency resource of most mobile communication operators has been severely fragmented. To deal with it, 3GPP also proposes non-contiguous CA in LTE-A system to improve spectrum efficiency. CCs in intraband non-contiguous CA are in the same frequency band, but frequency intervals exist between them.
3) Interband non-contiguous CA supports the aggregation of non-adjacent CCs in

the different bandwidth. It is obvious that interband non-contiguous CA takes the best use of frequency fragments in the overall system, and can theoretically achieve a favorable performance. However, the implementation of the physical layer of interband CA is much more complex than intraband CA.

Fig. 2.1 shows an illustration of the three categories of CA technique.

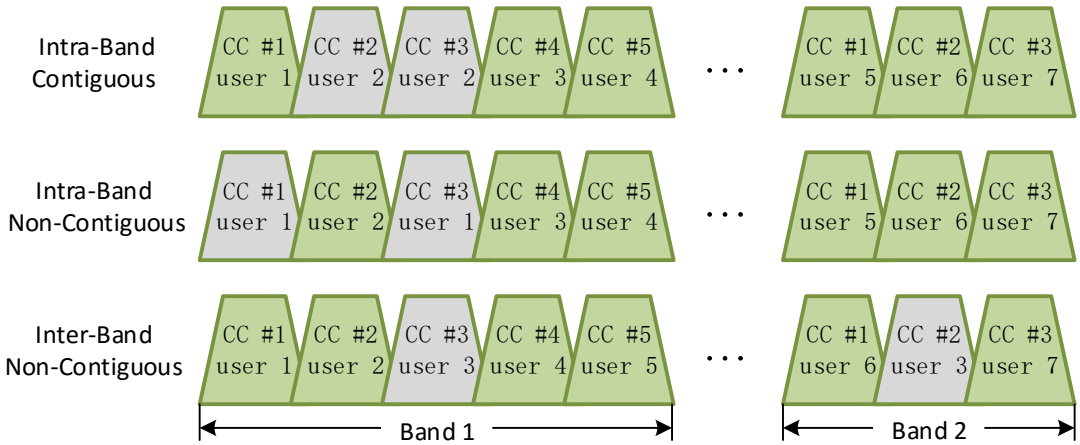


Fig. 2-1 Types of Carrier Aggregation [28]

Application of the CA technique brings new challenges, especially in terms of RRM of LTE-A systems. The first concern is the computation complexity of resource allocation. Apart from the RB allocation and power allocation involved in LTE system, scheduling CC resources for multiple UEs is also necessary, which brings serious difficulties for RRM. In LTE-A system, a UE is supposed to be equipped with multiple independent radio frequency chains (RFCs) to support CA. Consider a general LTE-A system consisting of M CCs and K UEs, and each UE is equipped with S RFCs. In this system, there are at most $\prod_{K=1}^{K}\binom{S}{M}$ possible results of CC allocation. The computation cost and allocation delay would exert heavy burden on the eNB, which in turn deteriorates the system performance, if an efficient allocation approach is absent. Second, unlike conventional resource allocation, CC allocation is generally performed before RB allocation and power allocation, which gives rise to great difficulties in evaluation on the quality of solution of CC allocation itself at the system level. Last but not least, with the limitation of available contiguous spectrum resources, interband non-contiguous CA scheme is more realistic to wireless operators. However, using non-contiguous CCs may introduce new constraints for resource allocation. For example, the limitation of UEs' capability of transmitting data on multiple CCs in the same time restricts the achievable performance of CA.

2.2 Heterogeneous Network

Dense network is a new cellular network deployment where a large number of low-power and low-cost access points are employed inside of the existing macro cell coverage. In dense networks, access points can be deployed flexibly in indoor or outdoor areas where channel quality is unfavorable, therefore avoiding the influence of obstacles on radio signal transmissions. At the same time, the introduction of access point shortens the transmission distance between BSs and users, reduces the path loss in the process of the wireless signal transmission, and results in improvement of the quality of the signal transmission [29][30] [31] [32] [33] [34].

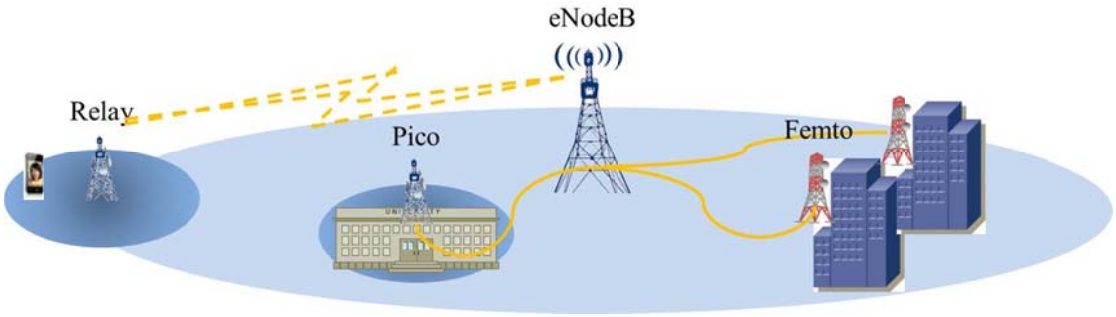


Fig. 2-2 HetNet

3GPP introduces a dense deployment, defined as Heterogeneous Network (HetNet), into LTE-A systems. Fig. 2-2 shows a simple example of a HetNet. Low-power access points in HetNet are sometimes referred to as small cell sites [33][34]. According to the functions, small cell sites can be classified into three categories [35], they are Pico, Femto and Relay. Compared to enhanced NodeBs (eNBs), which are macro BSs in LTE-A systems, these access points have characteristics of smaller transmission power, more flexible deployment and lower cost. In general, Pico [36][37] is configured with an omni-directional antenna without sectoring. Its transmission power is in the range between 250 mW and 2 W. The Pico is usually used for outdoor transmissions, but it also can be used for indoor transmission (at this time the transmission power no more than 100 mW) if necessary. Since it can be considered as an eNB with lower transmission power from the perspective of network architecture, Pico also uses X2-based backhaul as standardized in LTE-A systems. Pico is usually deployed according to the pre plan, so it's the location is relatively fixed. Pico is suitable for the crowded scenarios, such as shopping centers, activity centers, and so on, since it is beneficial for accommodating

more users to access the cellular network at the same time, as well as improving transmission quality.

Femto [38][39] is majorly used for indoor transmission, and its transmission power is no more than 100 mW. Femto is connected to the eNB with Digital Subcarrier Line (DSL) or cable. According to whether the user access to Femto is restricted, Femto cells can be divided into two categories: open Femto and closed Femto. From the user's point of view, open Femto is similar to Pico in features, since the connection between user and open Femto can be established a connection at any time when the user needs. However, closed Femto allows only authorized users to access while restricts all the external users. It is obvious that close Femto is capable to improve the safety and quality of transmission for authorized users. However, since omni-direction antenna is used, closed Femto is a potential interference source to the external radio transmissions in the neighborhood.

Different from Pico and Femto, the backhaul of a Relay [40][41] is usually a wireless link. It is referred to as an in-band Relay if the frequency band used for backhaul is the same to that for uplink and downlink transmission; otherwise, it is an out-band Relay. Obviously, physical techniques of out-band Relay are less challenged since there is no interference between backhaul transmissions and desired data transmissions. However, the spectrum efficiency is decreased because it is necessary to assign dedicated frequency band for backhaul. As a result, 3GPP prefers in-band Relay for the future mobile communication systems.

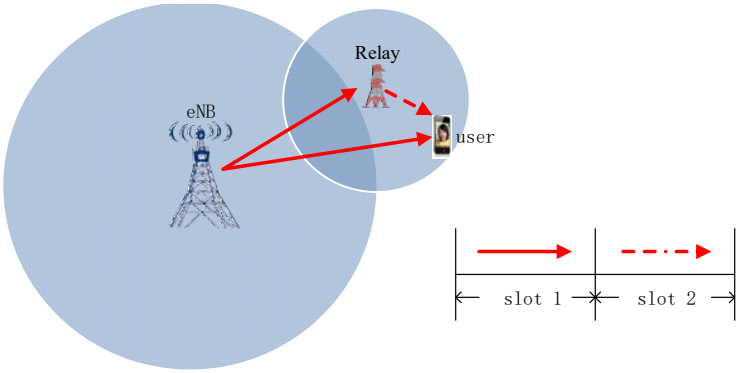


Figure 2-3 Illustration of transmission through relay

In Relay systems, a transmission lasts two time slots. Take a downlink transmission as an example. The desired data is transmitted by the eNB to both the user and the Relay in the first slot (slot 1 in Fig. 2-3); and then in the next time slot (slot 2 in Fig. 2-3), Relay transmits the processed data to the user. The example is

illustrated in Fig. 2-3. As a result, the user receives two independent versions of the desired data. The error probability of the transmission is reduced by the proper combination processing at the receiver. This is also known as a diversity gain obtained by introducing Relays into transmissions. In the scenario where it is impossible to realize Lines-of-Sight (LOS) transmissions between the eNB and users, Relays are also beneficial for covering holes existing in the coverage area.

Relay nodes can work in half-duplex mode or full-duplex mode. This kind of classification is majorly used for in-band Relays. As to out-band Relays, signals that are being transmitted by Relays will not interfere with the signals that are being received by the Relays, since different frequency bands are used. However, interference occurs in in-band Relay systems. It is reasonable to make Relay systems work in half-duplex mode, which means receiving and transmitting at Relays will not happen at the same time, in order to avoid the interference. If full-duplex mode is utilized, it is possible to spatially divide the receiving processing from the transmitting processing by adjusting the locations and angles of antennas. Generally, a Relay is equipped with a directional antenna pointing to the eNB, and an omni-directional antenna for communicating with users.

Table 2-1 gives a comparison of the main application scenarios and technical features of eNB, Pico, Femto and Relay.

Table 2-1：Application and characteristic of TPs in HetNets

Type	Applications	Transmission power	Radius of coverage	backhaul
eNB	Macro coverage	5~40 W	0.5~2 km	Dedicated links, X2
Pico	Crowded area outside	250 mW~2 W	≤300 m	Dedicated links, X2
Femto	Office and home	≤100 mW	≤50 m	DSL、cables
Relay	Holes of macro coverage	250 mW~2 W	≤300 m	Wireless links，X2

2.3 Coordinated Multiple Points

HetNet is a promising deployment for mobile communications systems. However, HetNet suffers from severe intercell interference, which seriously affects its performance. Compared to traditional cellular network, intercell interference is exacerbated in dense-deployed HetNet due to the shorter distance between BSs to

users. In one hand, the number of interfering sources in the neighborhood is increased since distances between BSs are shorter than before; on the other hand, interfering signals are strengthened since the distance between interference BS and the transmitting user is also shortened. Intercell interference affects the quality of transmissions to the user in cell-edged areas, resulting in unfavorable performance of the overall network. Therefore, avoiding intercell interference is one of the major subjects in the field of mobile communication systems.

Frequency Reuse Partitioning (FRP) [42] is one of the typical techniques for combating intercell interference. It splits the bandwidth of a system into several independent parts, and uses each part in different areas to avoid intercell interference. The benefit of interference avoidance brought by FRP is at the cost of spectrum efficiency of the system. Although several improved solutions are proposed to enhance spectrum efficiency, such as Partial Frequency Reuse (PFR) [43], Soft Frequency Reuse (SFR) [44], FRP technique is still not capable to meet the demands of future mobile communication systems.

Different from FRP, CoMP reduces intercell interference by allowing coordination and cooperation between adjacent BSs with the universal frequency reuse of the network, so that favorable data rates of cell-edged users and spectrum efficiency can be obtained [45][46][47]. CoMP technique can be considered as a large-scaled Multiple-Input Multiple Output (MIMO) system composed of several BSs. As in MIMO system, the desired signal of a downlink transmission with CoMP is strengthened by contributing multiple transmission antennas into the direction pointing to the target user. For this reason, CoMP technique is also called as network MIMO [48] sometimes. In order to realize cooperation between BSs, it is necessary to connect BSs via backhaul links to exchange control information and data between BSs. In addition, a Control Unit (CU) is also needed as a command center. The CU can be an individual entity or an entity embedded in a BS. It collects channel information from BSs that are under its control, and jointly schedules radio resource according to given algorithms. The scheduling results will be sent by the CU to each BS via backhaul links. At last, each BS performs transmissions with the resource assigned by the CU.

According to the different ways to cooperate, CoMP techniques are classified into two categories: coordinated scheduling/coordinated beamforming (CS/CB) and joint processing (JP) [49]. As illustrated in Fig. 2-4, CS/CB CoMP makes neighboring cells to jointly precoding according to global channel state information (CSI) to avoid potential

interference. In this case, user's data is transmitted by its home BS (i.e., the BS that the user has been registered with) only, so that no data needs to be shared between BSs via backhaul links. In essence, a multicell scenario with CS/CB CoMP can be modeled as an interference channel in Information Theory. However, due to the complexity of the interference channel and the cellular network, the Capacity Region of a CS/CB CoMP has not yet been conclusive. When CS/CB CoMP is employed, several BSs jointly precode symbols that are going to be transmitted to a specific user according to the known CSI between each BS and the user, aiming at reducing intercell interference. One of the typical precoding methods is Dirty Paper Coding (DPC), which is proved to be the optimal solution [50]. DPC is nonlinear and difficult to be used in large-scale systems due to the high-level computational complexity. Therefore, it is usually as a benchmark to evaluate performance of precoding design. Linear precoding is capable to approximate the performance of DPC with much lower computation, which makes it more valuable for practical applications. In this way, linear precoding of CS/CB CoMP attracts many attentions in both academic and industrial areas. Effective linear precode can be obtained based on the channel matrix (each elements of which represents the channel coefficient between a transmit antenna of a BS and a receive antenna of the user). The common solutions include QR decomposing [51], SVD decomposing [52], figuring out the pseudo inverse matrix [53] and so on. However, these solutions have a common restriction that the number of receiving antennas must be no less than that of transmitting antennas in order to ensure the operations of the channel matrices are reasonable. This means that the number of users that can be served at the same time is very limited.

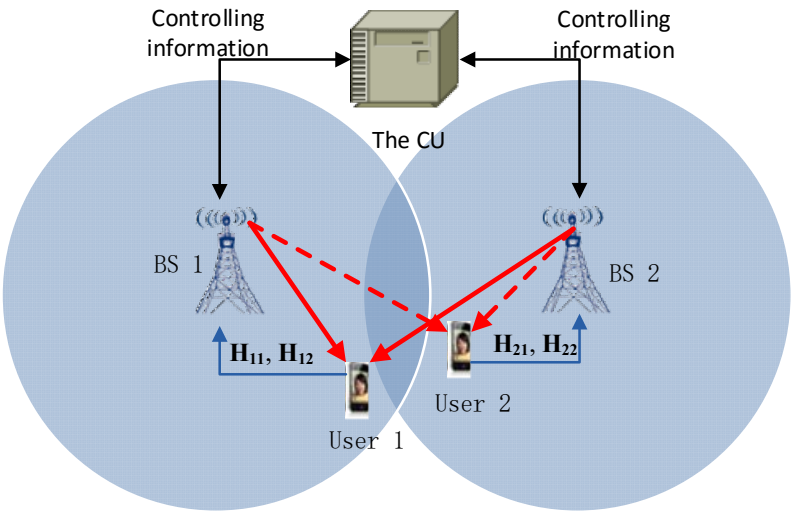
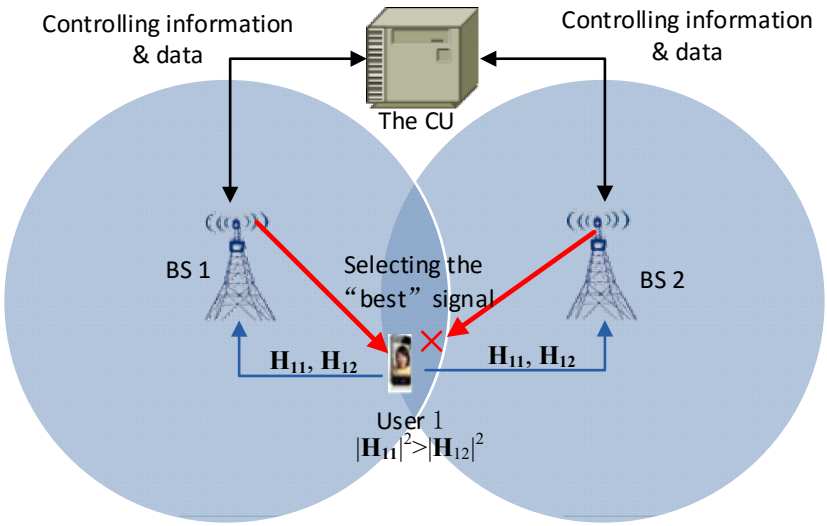
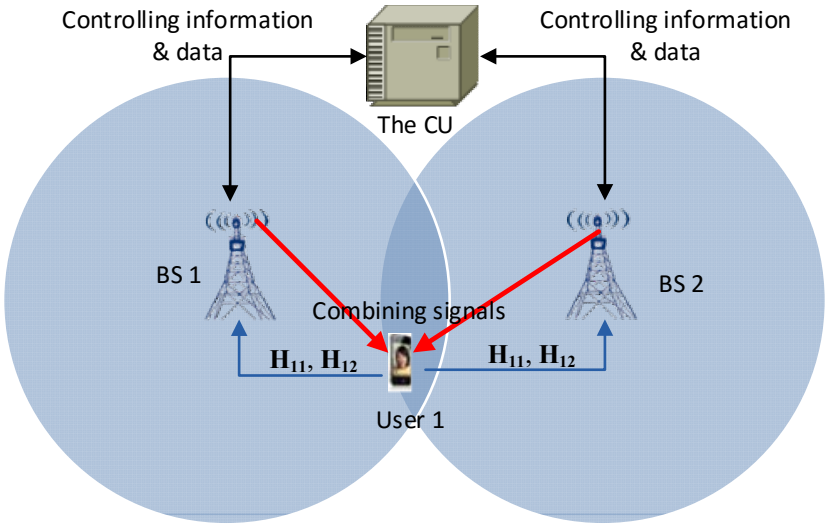


Fig. 2-4 Illustration of CS/CB CoMP



(a) DCS CoMP



b) JT CoMP

Fig. 2-5 Illustrations of JP CoMP

Different from CS/CB CoMP, JP CoMP takes use of intercell interference instead of avoiding it. JP CoMP allows one or more BSs in the neighborhood to transmit the same data to a specific user at the same time, as shown in Fig. 2-5. In this way, not only can the interference be decreased, but also can the desired signal be strengthened. According to the different transmission modes, JP CoMP is further divided into two categories: Dynamic Cell Selection (DCS) [54][55] and Joint Transmission (JT) [56][57]. When DCS CoMP is employed, the BS with the best channel condition to the specific user at present will be selected to transmit dataas shown in Fig 2-5(a). It is noticeable that the selected BS will not always be the home BS of the user due to the time variance of wireless channels. Compared to non-CoMP

transmission, the strength of the desired signal is improved thanks to the dynamic selection of transmitting BS according to instantaneous channel conditions in DCS CoMP transmission. To avoid intercell interference and further improve the quality of transmission, BSs that are not selected to transmit data in a DCS CoMP transmission can maintain mute at the same time and frequency resource. In any way, DCS CoMP is helpful for cell-edged users to obtain better quality of transmission.

Regardless of computational complexity, JT CoMP can achieve much better performance than the others. JT CoMP allows a cluster of BSs to transmit data signal to a specific user at the same time and frequency resource, as shown in Fig 2-5(b). Due to the independence of each channel between BSs and the user, the user can obtain several independent versions of the desired data, which brings significantly improvement to the signal strength at receiver after combination. At the same time, intercell interference in the considered cluster is completely avoided since all the BSs in it are source nodes of desired data in this case.

It should be noted that, as mentioned in the [48], CoMP technology is more suitable for the transmission of cell edge users. A cell edge user usually refers to a user located at the edge of the coverage area of its own home base station, and is close to the coverage area of one or more neighboring BSs. The channel gains between the users and the neighboring BSs are at same level of the channel gain to its home BS. In this case, the influence caused by intercell interference is so severe that transmissions to the user could probably be unsuccessful. CoMP technology takes use of these interference sources, so that the adjacent multiple BSs jointly transmit data signals and spacial diversity gain can be obtained to improve the quality of transmissions. Instead, as to users located closed to their home BSs, CoMP brings limited benefits to the transmissions at non-ignorable cost (overhead, processing computing, etc.). In this point of view, it is better to use CoMP technology for cell edge users only, instead of all the users in the network.

2.4 Conclusion

This chapter gives an introduction of three promising technologies proposed in LTE-A systems, including CA, HetNet and CoMP. CA is able to flexibly utilize limited spectrum efficiency, and significantly improve peak data rates in this way. HetNet is a dense deployment of cellular networks. By adding low-power access points into a macro cell,

HetNet can dramatically increase the throughput of the network. However, there are several challenges to HetNet. The major one of them is intercell interference that leads to failure transmissions of cell edge users, and therefore decreases the performance of the network. To combat with intercell interference, CoMP technology is proposed by 3GPP. CoMP mitigates the effects of intercell interference by allowing several BSs to cooperate with each other. The three technologies have attracted many attentions in both academic and industrial area.

Chapter 3 Resource allocation in LTE-A system with Carrier Aggregation

CA technique is proved to be beneficial for improving spectrum efficiency and therefore enhance network capacity. However, the use of CA brings several new challenges, especially in the term of radio resource allocation. In this chapter, we investigate the resource allocation problem in LTE-A systems, in order to maximize the gains from CA technique. First, we formulate the resource allocation as an optimization problem, aiming at maximizing sum capacity with multiple practical constraints. Then, the formulated problem is divided into three-stage of CC, RB and power allocation, in order to reduce the computational complexity. Specifically, for the CC allocation, we propose a Cross Entropy (CE) based greedy algorithm, with an approximate estimation of RB allocation and the equal power distribution. Given the CC allocation results, we design a RB exchange-based allocation algorithm to further improve the sum capacity. At last, we propose a Particle Swarm Optimization (PSO) based Power Allocation (PA) scheme. Extensive numerical simulations are launched to verify the efficiency of the proposed algorithms.

3.1 System Model

With CA, a user equipment (UE) is supposed to be equipped with multiple independent RFCs and can be simultaneously scheduled on multiple CCs. CC allocation determines which CC can be used by each RFC of users. On the base of CC allocation, RB and power allocation can be implemented with less computational complexity. The procedure of RRM in a mobile system with CA is illustrated by Fig. 3-1.

The goal of CC allocation is to assign each coming UE onto certain CC(s). In this paper, we present an estimation approach of RB-allocation to assist CC allocation and an average power distribution on per RB is adopted. Built on the approximations, we design a novel CE based greedy algorithm for CC allocation (CEGA-CCA). Given the solution of CC-allocation, RB-allocation can be further divided into the two categories: Independent scheduling in each CC (IC scheduling) and across CC scheduling (AC scheduling). IC scheme refers to allocating RBs in each CCs without considering others. The advantage of IC scheme is its compatibility with existing RB allocation algorithms

in LTE system. However, IC scheme is obviously a suboptimal solution. On the other hand, AC scheme takes all CCs into consideration when allocating RBs, and it hereby can achieve an overall better resource allocation than that of IC scheme. The main drawback of AC scheme lies in the higher complexity. To solve the RB allocation efficiently, this chapter studies the problem under the given CC allocation, and designs a modified greedy algorithm based on a RB exchange strategy. Finally, to further balance the tradeoff between fairness and sum capacity, under the condition of the given CC allocation and the RB-allocation results before, we also investigate the power allocation by using PSO algorithm, a kind of heuristic method, to make this work complete. Fig. 3-1 illustrates the overall resource allocation strategy.

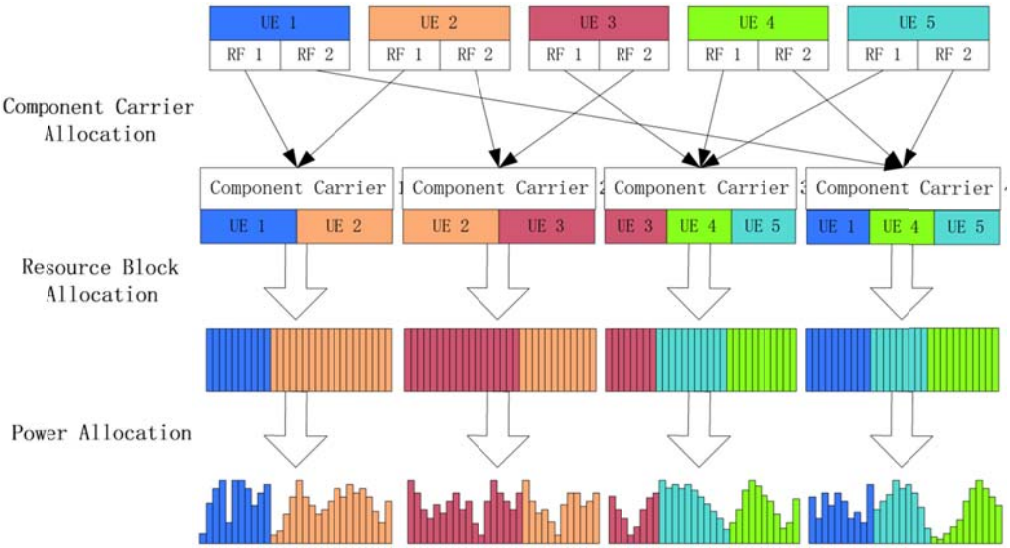


Figure 3-1 Illustration of the overall resource allocation strategy

In this chapter, we consider a LTE-A system comprising M CCs and K UEs. The CC m consists of N_m OFDM RBs, and each UE is equipped with S RFCs. Denote RFC index of UE k by $\psi_{k,s}^{m} \in \{0,1\}$, where $\psi_{k,s}^{m} = 1$ means RFC s of UE k is allowed to transmit data over CC m, and means $\psi_{k,s}^{m} = 0$ otherwise. Similarly, denote RB index by $\beta_k^{m,n} \in \{0,1\}$, where $\beta_k^{m,n} = 1$ means RB n of CC m is allocated to UE k. Let $p^{m,n}$ denotes the transmit power on RB n of CC m, then the data rate of RFC s of UE k on CC m can be expressed as

$$R_{k,s}^{m} = \psi_{k,s}^{m} \left(\sum_{n=1}^{N_m} \beta_k^{m,n} W \log_2 \left(1 + \frac{p^{m,n} \left(h_{k,s}^{m,n} \right)^2}{\Gamma_k N_0 W} \right) \right) \tag{3-1}$$

where N_0 is the two sided spectral density of additive white Gaussian noise (AWGN),

W is the bandwidth of each RB, $\Gamma_k = \frac{1}{3}Q^{-1}\left(\frac{\mathrm{BER}_k}{4}\right)^2$ is the signal-noise ratio (SNR) gap due to modulation [58], and BER_k is the requirement of bit error rate (BER) for UEs when M-QAM is adopted, $h_{k,s}^{m,n}$ is the channel gain for RFC s of UE k on RB n of CC m due to pass loss, shadowing and fading. For simplification, we assume that RFCs are independent and identical for the same UE, and therefore $h_{k,s}^{m,n}=h_k^{m,n}$ and $\psi_{k,s}^m = \psi_k^m$. Then the sum capacity of UE k can be expressed as

$$R_k = \sum_{m=1}^{M} R_k^m = \sum_{m=1}^{M}\sum_{s=1}^{S} R_{k,s}^m . \tag{3-2}$$

Fairness is an important performance metric in multiuser wireless communication systems [59] [60]. By exploiting user fairness, we can explicitly control the capacity ratios among UEs, and ensure that each UE, especially for cell-edge UEs, achieves the expected data rate. In this paper, we adopt Jain's index to evaluate fairness the same as in [59],

$$F := \frac{\left(\sum_{k=1}^{K} R_k\right)^2}{K\sum_{k=1}^{K} R_k^2} . \tag{3-3}$$

The maximum value of $F = 1$ will be achieved when $R_1 = R_k$, $\forall k$. In order to make sure that every UE can be scheduled in each allocation, we define a non-continuous point $F = 0$, if $\exists k$ whose $R_k = 0$.

Table 3-1 Important Denotations

Denotations	**Meaning**
M	number of CCs
K	number of UEs
N_m	number of RBs in CC m
S	Number of RFCs of each UE
$\psi_{k,s}^m$	RFC index
$\beta_k^{m,n}$	RB scheduling index
$p^{m,n}$	transmit power on RB n of CC m
$R_{k,s}^m$	data rate of RFC s of UE k on CC m
R_k	the sum capacity of UE k

The goal of this paper is to find the optimal resource allocation to maximize system sum capacity while guaranteeing fairness among UEs. The system capacity is the optimal object, and the fairness is constraints. Mathematically, the problem can be formulated as a constrained optimization problem as,

$$
\begin{aligned}
&(\boldsymbol{\psi},\boldsymbol{\beta},\mathbf{p}) \in \arg\max \sum_{k=1}^{K} R_k \\
\text{s.t.}\quad &\text{C1}: \quad \sum_{m=1}^{M} \psi_k^m \le S, \psi_k^m \in \{0,1\}, \forall k \\
&\text{C2}: \quad \sum_{k=1}^{K}\sum_{n=1}^{N_m} \beta_k^{m,n} = 1, \beta_k^{m,n} \in \{0,1\}, \forall m \\
&\text{C3}: \quad \sum_{n=1}^{N_m} p^{m,n} \le P_{\text{total}}^m, p^{m,n} \in \left[0, P_{\text{total}}^m\right], \forall m,n \\
&\text{C4}: \quad F = 1
\end{aligned}
\qquad , \tag{3-4}
$$

where $\boldsymbol{\psi} := \left[\psi_k^m\right]_{K\times M}$ and $\boldsymbol{\beta} := \left[\beta_k^{m,n}\right]_{K\times M\times(\max N_m)}$ indicate the assignments of CC and RB, respectively. The constraint C1 implies that one UE can transmit data over *S* CCs at most; C2 indicates that each RB can only be assigned to one UE at a time; C3 means the total power of all RBs should be positive but less than the power limitation P_{total}^m of CC *m*, $\forall m$; and C4 is the fairness constraint.

The initial problem in Eq. (3-4) is a typical mix-integer nonlinear programming (MINP) problem, which is believed difficult to be solved in polynomial time. Here, we adopt a skill of transferring objective function that combines the constraint C4 into the original objection function. The rationale behind this transformation is from the following proposition.

Proposition 1: Under the constraint condition of $F = 1$, if $(\boldsymbol{\psi},\boldsymbol{\beta},\mathbf{p}) \in \arg\max \sum_{k=1}^{K} R_k$, then $(\boldsymbol{\psi},\boldsymbol{\beta},\mathbf{p}) \in \arg\max R_k$.

Proof: Suppose $\exists(\boldsymbol{\psi}', \boldsymbol{\beta}', \mathbf{p}')$, such that $\min R_k' > \min R_k$. According to definition in Eq. (3-3) and the constraint $F = 1$, we obtain $\sum_{k=1}^{K} R_k' = K \min R_k' > \sum_{k=1}^{K} R_k = K \min R_k$, which is contradictory with $(\boldsymbol{\psi},\boldsymbol{\beta},\mathbf{p}) \in \arg\max \sum_{k=1}^{K} R_k$, thus $(\boldsymbol{\psi}', \boldsymbol{\beta}', \mathbf{p}')$ does not exist. ▌

According to **Proposition 1**, the initial problem can be transformed into the following one,

$$
\begin{aligned}
&(\mathcal{P}) \quad (\boldsymbol{\psi},\boldsymbol{\beta},\mathbf{p}) \in \arg\max\min R_k \\
\text{s.t.} \quad &\text{C1}: \quad \sum_{m=1}^{M} \psi_k^m \le S, \psi_k^m \in \{0,1\}, \forall k \\
&\text{C2}: \quad \sum_{k=1}^{K}\sum_{n=1}^{N_m} \beta_k^{m,n} = 1, \beta_k^{m,n} \in \{0,1\}, \forall m \\
&\text{C3}: \quad \sum_{n=1}^{N_m} p^{m,n} \le P_{\text{total}}^m, p^{m,n} \in \left[0, P_{\text{total}}^m\right], \forall m,n
\end{aligned}
\tag{3-5}
$$

To simplify problem ($\mathcal{P}$), we adopt a decomposed framework that decomposes the problem into three suboptimal problems, CC allocation, RB allocation and power allocation. Specifically, for CC allocation, we first present an effective method to approximate RB allocation and power allocation, and then solve the CC allocation problem under those approximations. Similarly, we solve RB allocation under the given CC allocation results and approximate power-allocation. When CC allocation and RB-allocation are given, the problem of power-allocation can be derived naturally. The decomposed suboptimal problems can be written as follows:

$$
\begin{aligned}
&(\mathcal{P}1) \quad \boldsymbol{\psi} \in \arg\max\min R_k(\boldsymbol{\beta},\mathbf{p}) \\
\text{s.t.} \quad &\text{C1} \qquad \sum_{m=1}^{M} \psi_k^m \le S, \psi_k^m \in \{0,1\}, \forall k
\end{aligned}
\tag{3-6}
$$

$$
\begin{aligned}
&(\mathcal{P}2) \quad \boldsymbol{\beta} \in \arg\max\min R_k(\boldsymbol{\psi},\mathbf{p}) \\
\text{s.t.} \quad &\text{C2} \qquad \sum_{k=1}^{K}\sum_{n=1}^{N_m} \beta_k^{m,n} = 1, \beta_k^{m,n} \in \{0,1\}, \forall m
\end{aligned}
\tag{3-7}
$$

$$
\begin{aligned}
&(\mathcal{P}3) \quad \mathbf{p} \in \arg\max\min R_k(\boldsymbol{\psi},\boldsymbol{\beta}) \\
\text{s.t.} \quad &\text{C3} \qquad \sum_{n=1}^{N_m} p^{m,n} \le P_{\text{total}}^m, p^{m,n} \in \left[0, P_{\text{total}}^m\right], \forall m,n
\end{aligned}
\tag{3-8}
$$

The following sections present the respective solution to ($\mathcal{P}1$), ($\mathcal{P}2$) and ($\mathcal{P}3$). The entire strategy of RRM in a considered CA system is demonstrated in Fig. 3-2.

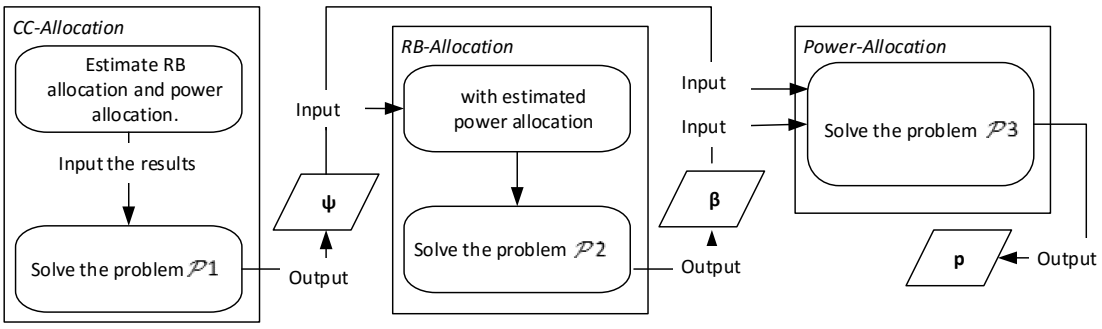


Figure 3-2 Flowchart of the entire strategy for RRM

3.2 Component Carriers Allocation

In this section, we first give an approximate estimation of **β** and **p**, respectively. Under these estimations, we present a Cross Entropy based greedy algorithm for CC allocation algorithm (CEGA-CCA).

A. Estimation of RB-Allocation and Power Allocation

For power allocation **p**, we assume that power is equal distributed on all RBs of the same CC. Define the virtual power $P := \frac{P_{\text{total}}^1}{N_1}$, and then we can get

$$p^{m,n} = p^{m,-} = \frac{P_{\text{total}}^m}{N_m} = \frac{P_{\text{total}}^m N_1}{P_{\text{total}}^1 N_m} P, \forall m,n \tag{3-9}$$

Define

$$H_k^{m,n} = \frac{\left|h_k^{m,n}\right|^2 p^{m,-}}{\Gamma N_0 W P} \tag{3-10}$$

According to the channel fading theory, large-scale fading is the domination of $H_k^{m,n}$. Hence, for a specific UE *k*, all $H_k^{m,n}$ over the bandwidth are supposed to fluctuate in the same level. We can estimate the sum capacity of UE *k* in CC *m* as follows, based on the approximation $\ln(1+x)-\ln(1+x') \approx (x'-x)\frac{1}{1+x'}, |x-x'| < \varepsilon$:

$$\begin{aligned} R_k^m &= \sum_{n=1}^{N_m} \beta_k^{m,n} W \log_2\left(1+PH_k^{m,n}\right) = \sum_{n=1}^{N_m} \frac{\beta_k^{m,n} W}{\ln 2} \ln\left(1+PH_k^{m,n}\right) \\ &\approx \frac{1}{\ln 2} \sum_{n=1}^{N_m} \beta_k^{m,n} W \left(\left(PH_k^{m,n} - P\overline{H_k^m}\right)\tan\theta + \ln\left(1+P\overline{H_k^m}\right)\right) \\ &= \frac{W}{\ln 2}\left(0 + \left|\Omega_k^m\right| \ln\left(1+P\overline{H_k^m}\right)\right) = \left|\Omega_k^m\right| W \log_2\left(1+P\overline{H_k^m}\right) \end{aligned} \tag{3-10}$$

where Ω_k^m is the set of RBs assigned to UE *k* over CC *m*, and operator |·| is the cardinality number of the set which satisfies

$$\left|\Omega_k^m\right| = \sum_{n=1}^{N_m} \beta_k^{m,n} \tag{3-11}$$

$$\overline{H_k^m} = \sum_{n=1}^{N_m} \beta_k^{m,n} H_k^{m,n} \Big/ \left|\Omega_k^m\right| \tag{3-12}$$

and

$$\tan\theta = \frac{d\left(\ln\left(1+P\overline{H_k^m}\right)\right)}{d\left(1+P\overline{H_k^m}\right)} = \frac{1}{\left(1+P\overline{H_k^m}\right)} \tag{3-13}$$

According to the constraint of fairness, we can get

$$\sum_{m=1}^{M}\psi_k^m\left|\Omega_k^m\right|\log_2\left(1+P\overline{H_k^m}\right) = \sum_{m=1}^{M}\psi_{k'}^m\left|\Omega_{k'}^m\right|\log_2\left(1+P\overline{H_{k'}^m}\right) \tag{3-14}$$

However, it is nontrivial to get the solution $\boldsymbol{\psi}$ because we cannot obtain the accurate value of $\left|\Omega_k^m\right|$ prior to RB allocation. Hence, an approximate approach is used that we assume the fairness is satisfied in each CC, i.e.

$$\left|\Omega_k^m\right|\log_2\left(1+P\overline{H_k^m}\right) = \left|\Omega_{k'}^m\right|\log_2\left(1+P\overline{H_{k'}^m}\right), \forall m \tag{3-15}$$

and

$$\sum\nolimits_{k=1}^{K}\left|\Omega_k^m\right| = N_m, \forall m \tag{3-16}$$

Under the assumption that equal power is used for each transmission, the approximate number of RBs assigned to UE *k* can be obtained by jointly solving Eq. (3-15) and Eq. (3-16).

$$\left|\Omega_k^m\right| \approx \left\lceil \frac{\psi_k^m N_m}{\log_2\left(1+P\overline{H_k^m}\right)\sum_{k'=1}^{K}\left(\psi_{k'}^m \Big/ \log_2\left(1+P\overline{H_{k'}^m}\right)\right)} \right\rceil, \forall m,k \tag{3-17}$$

The estimated sum capacity of UE *k* over CC *m* can be written as follows

$$R_k^{m(\text{ets})} = \frac{N_m W \psi_k^m}{\sum_{k'=1}^{K}\left(\psi_{k'}^m \Big/ \log_2\left(1+P\overline{H_{k'}^m}\right)\right)}, \forall m,k \tag{3-18}$$

and therefore, $R_k^{(\text{ets})} = \sum_{m=1}^{M} R_k^{m(\text{ets})}$.

B. **GA-CCA: Greedy Algorithm for CC allocation.**

Greedy algorithm is a typical method to achieve the objective. In this section, we first present a greedy algorithm for CC allocation (GA-CCA). In each step, GA-CCA aims at improving the capacity of the UE who has the minimal capacity (named a minimal capacity UE). The proposed GA-CCA algorithm follows the three steps as shown in Fig. 3-3.

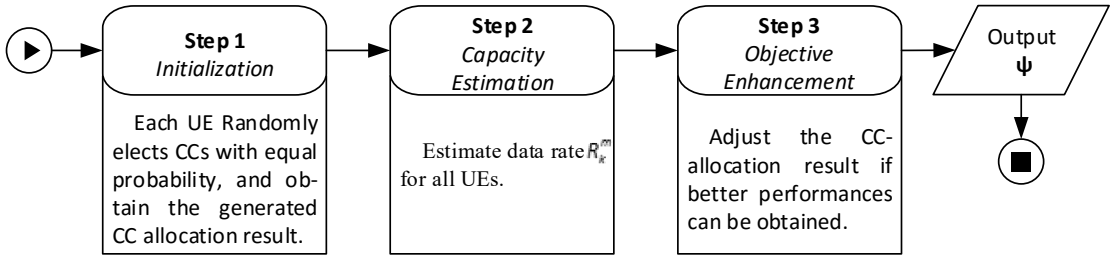


Figure 3-3 Flowchart of GA-CCA

The details of the proposed GA-CCA algorithm are performed as the following steps:

STEP 1 (Mobile Hashing Allocation): At the beginning of CC allocation, UEs select CCs with equal probability;

STEP 2 (Capacity Estimation): Estimate capacity R_k^m for all UEs according to (3-18);

STEP 3 (Objective Enhancement): Find the minimal capacity UE $k_{\min}$, i.e.,

$$k_{\min} \in \arg\min_{1\le k\le K} R_k \tag{3-19}$$

Find CC m^* where UE $k_{\min}$ has the minimal capacity if $\psi_{k_{\min}}^m \neq 0$, i.e.,

$$m^* \in \arg\min_{1\le m\le M}\{R_{k_{\min}}^m\} \tag{3-20}$$

If CC m^* satisfies Eq. (3-20), it may not suit UE $k_{\min}$ or be over-crowded. The proposed algorithm attempts at eliminating such allocations. Therefore, find CC μ if it exists, such that

$$\mu \in \arg\max_{1\le \mu\le M, \mu\neq m^*} \sum_m R_{k_{\min}}^m \tag{3-21}$$

and set $\psi_{k_{\min}}^{m^*} = 0$ and $\psi_{k_{\min}}^{\mu} = 1$.

This process is repeated until we cannot find μ in (3-21), and the search process ends. The detailed implementation of GA-CCA is summarized as in **Algorithm 3-1**.

Algorithm 3-1 Greedy CCA Algorithm

1: (Initialization): $\forall k, m$, set $\psi_{k_{\min}}^m = 0$ and $R_{k_{\min}}^m = 0$.

2: **for** $k \leftarrow 1$ to K **do**

3: Select S CCs from M CCs with equal probability.

4: **if** CC m is selected **then**

5: $\psi_{k_{\min}}^m = 1$

6: **end if**

7: **end for**

8: Find UE $k_{\min}$ that satisfies $k_{\min} \in \arg\min_k\{R_k\}$.

9: Find CC m^* that satisfies $m^* \in \arg\min_m\{R^m_{k_{\min}}\}$.

10: **While** $\exists\mu\in\{1,\ldots,M\}$ satisfies $\mu \in \arg\max_{\mu\neq m^*} \sum_m R^m_{k^*}$ **do**

11: Set $\psi^{m^*}_{k_{\min}} = 0$ and $\psi^{\mu}_{k_{\min}} = 1$.

12: Re-estimate capacity R^m_k for all UEs according to Eq. (3-18).

13: **end while**

C. **CEGA-CCA: Cross Entropy based GA-CCA.**

Although GA-CCA has low computation complexity, the quality of solution is tightly dependent on the random allocation stage and easy to trap into local extremum. In this section, we present a heuristic method to further improve the CC allocation on the basis of GA-CCA.

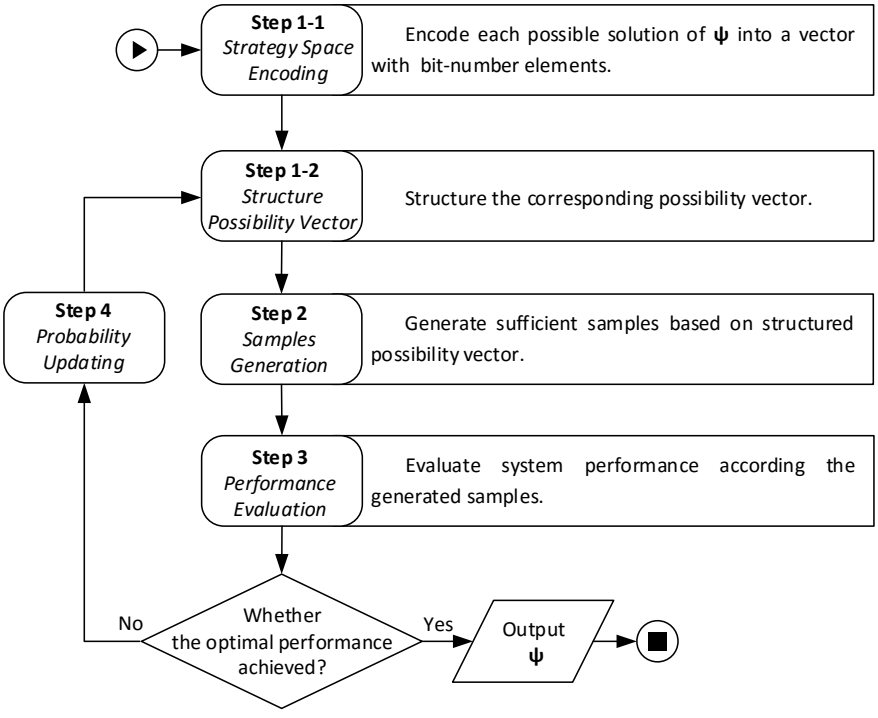


Figure 3-4 Flowchart of CEGA-CCA

CE is a novel heuristic method [61]. It is firstly used to simulate small probability matters, and then introduced to solve the optimization problem [62]. The basic steps of CE includes four steps: (1) Set up the solution space, in order to map the initial problem into the CE method; (2) Generate random samples from solution space according to some probability distribution; (3) Evaluate each generated samples according to the objective and constraints, and eliminate some inferior samples; (4)

Update the parameters of the probability distribution, in order to obtain better samples in the next iteration. A concrete understanding of the CE method is referred to [61]. The procedure of the proposed CE based GA-CCA (CEGA-CCA) method is illustrated in Fig. 3-4, and the details are explained to search for the optimal $\boldsymbol{\psi}$ in problem ($\mathcal{P}1$) as follows:

STEP 1 (Strategy Space Encoding). Set the iteration counter t=1. For UE k equipped with S RFCs and M CCs candidates, we can order its CC selection strategy lexicographically, and the strategy space is

$$\Theta_k = \left[\theta_k^1, \theta_k^2, ..., \theta_k^{\Lambda}\right] \tag{3-22}$$

where $\Lambda = \binom{M}{S}$ is the size of Θ_k, and

$$\begin{aligned} \theta_k^1 &= \left\{\psi_k^1 = 1, \cdots, \psi_k^S = 1, \psi_k^{S+1} = 0, \cdots, \psi_k^M = 0\right\} \\ \theta_k^2 &= \left\{\psi_k^1 = 1, \cdots, \psi_k^S = 0, \psi_k^{S+1} = 1, \cdots, \psi_k^M = 0\right\} \\ &\vdots \\ \theta_k^{\Lambda} &= \left\{\psi_k^1 = 1, \cdots, \psi_k^S = 0, \psi_k^{M-S+1} = 1, ..., \psi_k^M = 1\right\} \end{aligned} \tag{3-23}$$

Accordingly, we structure a Λ dimension probability vector in the t^{th} iteration $\mathbf{P}_{k,t}$ as:

$$\mathbf{P}_{k,t} = \left[P_{k,t}^1, P_{k,t}^2, ..., P_{k,t}^{\Lambda}\right], \sum_{i=1}^{\Lambda} P_{k,t}^i = 1 \tag{3-24}$$

where $p_{k,t}^i$ represents the probability that UE k select strategy θ_k^i.

Step 2 (Samples Generation). Use the probability vector $\mathbf{P}_{k,t}$ to generate random samples

$$\mathbf{X}_k(z) = \left[X_k^1(z), X_k^2(z), ..., X_k^{\Lambda}(z)\right], 1 \le z \le Z \tag{3-25}$$

where Z is the number of samples, $\mathbf{X}_k(z)$ is a Λ dimension vector with only one element "1" and (Λ-1) elements "0", and the probability of $X_k^i(z) = 1$ is $P_{k,t}^i$.

Step 3 (Performance Evaluation). Decode $\psi_k^m, \forall m$ according to the generated sample $\mathbf{X}_k(z)$, and then substitute the ψ_k^m into GA-CCA algorithm to calculate the value of objective function, denoted as V_z. Rearrange V_z in descending order as $V_1 \ge ... \ge V_z$. After that, let $\lambda = V_{\lceil Zq \rceil}$ be the sample q-quantile of the performance, where q denotes the quantile coefficient and $\lceil \bullet \rceil$ is the ceiling operation.

Step 4 (Probability Updating). Update the probability $P_{k,t}^{i}$ for all 1≤ k ≤K and 1≤ k ≤Λ as follows

$$P_{k,t}^{i}=\frac{\sum_{z=1}^{Z} I_{\{V_z\geq\lambda\}} X_k^i(z)}{\sum_{z=1}^{Z} I_{\{V_z\geq\lambda\}}} \tag{3-26}$$

where $I_{\{V_z\geq\lambda\}}$ is an indicated variable defined by

$$I_{\{V_z\geq\lambda\}}=\begin{cases}1 & \text{if} \quad V_z\geq\lambda\\ 0 & \text{otherwise}\end{cases} \tag{3-27}$$

STEP 5: Ends the iteration if a convergence criterion is satisfied, which might be a maximum number of iterations or $P_{k,t}^{i}$, ∀k, i; otherwise set t := t + 1 and go to **STEP 2**.

The implementation of CEGA-CCA is summarized as in **Algorithm 3-2**.

Algorithm 3-2 CE Enhanced Greedy CCA Algorithm

1: (Initialization): ∀k, m, set $\psi_{k_{\min}}^{m}=0$ and R_k^m=0.
2: for t←1 to T^{CE} do
3: for z←1 to Z^{CE} do
4: for k←1 to K do
5: Generate sample $\mathbf{X}_k(z)$, according to $\mathbf{P}_{k,t}$
6: for m←1 to M do
7: Calculate ψ_k^m according to $\mathbf{X}_k(z)$
8: end for
9: end for
10: V_z←GACCA Algorithm(**Step 2** and **Step 3**)
11: end for
12: Rearrange V_z in descending order as V_1 ≥…≥V_z
13: for k←1 to K do
14: for i←1 to Λ do
15: updating $P_{k,t}^{i}$ using Eq. (3-26)
16: end for
17: end for
18: end for

3.3 Resource Blocks Allocation

In this section, we investigate the RB-allocation problem under the given CC allocation results and estimated power allocation above. First, we use a greedy RB-allocation algorithm in [58] and [59]. Then, we present a RB exchange based greedy algorithm for RB allocation using the ratio of channel fading between two UEs in the same CC.

A. Greedy Resource Block Allocation Algorithm (MG-RBA)

The intuition behind greedy algorithm is to maximize the sum capacity by taking advantage of multiuser diversity as possible, while maintaining a coarse proportional fairness. Based on the greedy allocation, we design a modified greedy RB allocation algorithm (MG-RBA) to make it suited for our problem.

The proposed MG-RBA algorithm consists of two stages, as shown in Fig. 3-5. For the first stage, each UE selects one RB on the allocated CCs in sequence. After that, the minimal capacity UE has the priority to select the best one of rest RBs until all RBs are assigned to the corresponding UEs. The details of the proposed greedy RB allocation algorithm are summarized as in **Algorithm 3-3**. For clarity, we denote $\mathcal{K}_m$ and $\mathcal{A}_m$ as the set of RBs allocated to CC m and the rest RBs in CC m, respectively.

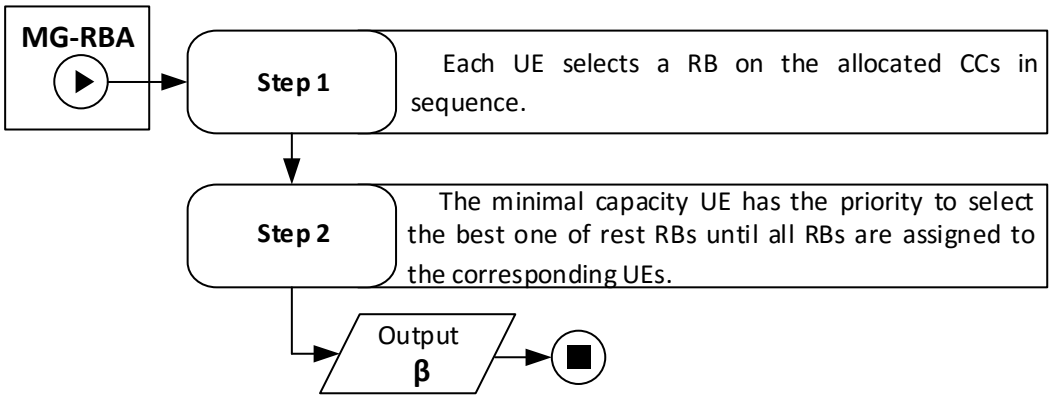


Fig. 3-5 Flowchart of MG-RBA

Algorithm 3-3 Greedy Resource Block Allocation Algorithm

Input: ψ , K, M, $|\Omega_k^m|$, N_m, $\rho=\emptyset$, $\mathcal{K}_m=\emptyset$, $\mathcal{A}_m=\emptyset$, $R_k=0$, $\forall m, k$

Output: RB-allocation results ρ.

1: **for** $m \leftarrow 1$ to M **do**

2: $\mathcal{A}_m$={1, …, N_m},

3: **if** $\psi_k^m = 1$ **then**

4: $\mathcal{K}_m$=$\mathcal{K}_m$+{k}

5: **end if**

6: **end for**

7: **for** $k \leftarrow 1$ to K **do**

8: **for** $m \leftarrow 1$ to M **do**

9: **if** $\psi_k^m = 1$ **then**

10: $n^* = \arg\max_n\{H_k^{m,n}\}$, $k \in \mathcal{K}_m$

11: ρ_k^{m,n^*}=1

12: R_k=R_k+W$\log_2\left(1+PH_k^{m,n^*}\right)$

13: $\mathcal{A}_m$=$\mathcal{A}_m$-{n^*}

14: $|\Omega_k^m|$=$|\Omega_k^m|$-1

15: **end if**

16: **end for**

17:**end for**

18:**for** $m \leftarrow 1$ to M **do**

19: **while** $\mathcal{A}_m \neq \emptyset$ **do**

20: Find $k_{\min}$ that satisfies $R_{k_{\min}} \leq R_k, \forall k \in K_m$

21: Find n^* that satisfies $H_{k_{\min}}^{m,n^*} > H_{k_{\min}}^{m,n}, \forall n$

22: $\rho_{k_{\min}}^{m,n^*}$=1,

23: $R_{k_{\min}}$=$R_{k_{\min}}$+W$\log_2\left(1+PH_{k_{\min}}^{m,n^*}\right)$

24: $\mathcal{A}_m$=$\mathcal{A}_m$-{n^*}

25: $\left|\Omega_{k_{\min}}^m\right|$=$\left|\Omega_{k_{\min}}^m\right|$-1

26: **if** $\left|\Omega_{k_{\min}}^m\right|$=0 **then**

27: $\mathcal{K}_m=\mathcal{K}_m-\{k_{\min}\}$

28: **end if**

29: **end while**

30:**end for**

RB Exchange based Resource Block Algorithm (RE-RBA).

Based on the proposed MG-RBA algorithm, we further proposed an RB exchange based algorithm (RE-RBA) to further balance the tradeoff between sum capacity and proportional fairness. Specifically, for given solution ρ of MG-RBA, we can obtain the capacity $R_k, \forall k$ under the equation power distribution assumption. After that, the algorithm finds the minimal capacity UE $k_{\min}$ and its maximal capacity CC $m_{\max}$. Then, we adopt a conservative strategy to exchange RB from some UE to UE $k_{\min}$ within CC m. According to [63], the candidate RB n^* transferring to $k_{\min}$ with lowest loss of sum capacity has to satisfy condition $\max_n\{H_{k_{\min}}^{m_{\max},n}/H_{k'}^{m_{\max},n}\}$, where $k'(k'\neq k_{\min})$ is the current owner of RB n^* in CC $m_{\max}$. If such RB exists, set $\rho_{k_{\min}}^{m_{\max},n^*}=1$, $\rho_{k'}^{m_{\max},n^*}=0$, and recalculate capacity $R_{k_{\min}}=R_{k_{\min}}+\text{Wlog}_2\left(1+PH_{k_{\min}}^{m_{\max},n^*}\right)$ and $R_{k'}=R_{k'}-\text{Wlog}_2\left(1+PH_{k'}^{m_{\max},n^*}\right)$. This process is repeated until we cannot find out such RB that can improve the capacity of certain $R_{k_{\min}}$, and the algorithm stops. Fig. 3-6 gives a flowchart of the proposed RE-RBA, and the detailed implementation of RE-RBA is summarized in **Algorithm 3-4**.

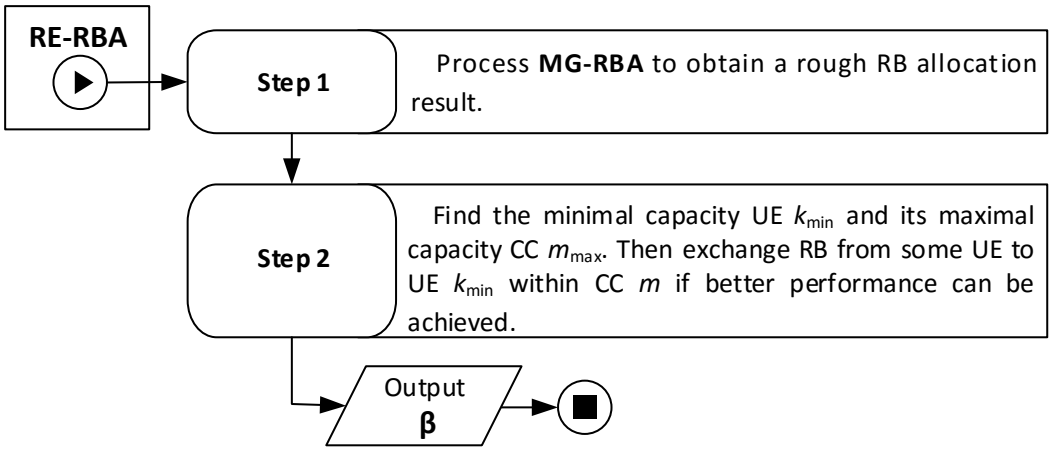


Figure 3-6 Flowchart of RE-RBA

Algorithm 3-4 RE-RBA Algorithm

1: (Initialization): $\boldsymbol{\psi}$, K, M, $|\Omega_k^m|$, N_m, $\rho=\emptyset$, $\mathcal{A}_m=\emptyset$, $R_k=0$, $\forall m$, k.

2: for $m \leftarrow 1$ to M do

3: $\mathcal{A}_m=\{1,\ldots, N_m\}$

4: if $\psi_k^m = 1$ then

5: $K_m=K_m+\{k\}$

6: end if

7: end for

8: for $k \leftarrow 1$ to K do

9: for $m \leftarrow 1$ to M do

10: if $\psi_k^m = 1$ then

11: $n^* = \arg\max_n\{H_k^{m,n}\}$, $k \in \mathcal{K}_m$

12: $\rho_k^{m,n^*} = 1$

13: $R_k = R_k + W\log_2\left(1+PH_k^{m,n^*}\right)$,

14: $\mathcal{A}_m = \mathcal{A}_m - \{n^*\}$

15: $|\Omega_k^m| = |\Omega_k^m| - 1$

16: end if

17: end for

18: end for

19: for $m \leftarrow 1$ to M do

20: while $\mathcal{A}_m \neq \emptyset$

21: Find $k_{\min}$ satisfying $R_{k_{\min}} \leq R_k$, $\forall k \in \mathcal{K}_m$.

22: For $k_{\min}$, find n^* satisfying $H_{k_{\min}}^{m,n^*} > H_{k_{\min}}^{m,n}$, $\forall n$

23: $\rho_k^{m,n^*} = 1$

24: $R_k = R_k + W\log_2\left(1+PH_k^{m,n^*}\right)$

25: $\mathcal{A}_m = \mathcal{A}_m - \{n^*\}$

26: $|\Omega_k^m| = |\Omega_k^m| - 1$

27: if $|\Omega_k^m| = 0$ then

28: $\mathcal{K}_m = \mathcal{K}_m - \{k_{min}\}$

29: end if

30: end while

31: end for

3.4 Power allocation

After RB-allocation, we can further improve the objective by distributing power on each RB. Substituting the obtained $\boldsymbol{\beta}$ with the given $\boldsymbol{\psi}$ into problem ($\mathcal{P}3$), the power allocation can be seen as a constrained continuous variable optimization. In this section, we propose a heuristic method based on the PSO algorithm. In contrast to other heuristic methods, such as genetic algorithm (GA) and ant colony optimization (ACO), PSO has better global searching capability at the beginning of the run and a local searching capability near the end of the run thanks to the connections among particles [64]. The standard PSO algorithm consists of the following steps: (1) Construct the particle to map the solution of interest problem; (2) Create the topology of particle swarm (particles' position); (3) Calculate fitness value for each particle and find out the best particle; (4) Updating particles' position; (5) Repeat step (3)-(4) until the stopping criterion is satisfied. Alone with the standard PSO algorithm, the procedure of the proposed CE based power allocation (PSO-PA) is illustrated in Fig. 3-6.

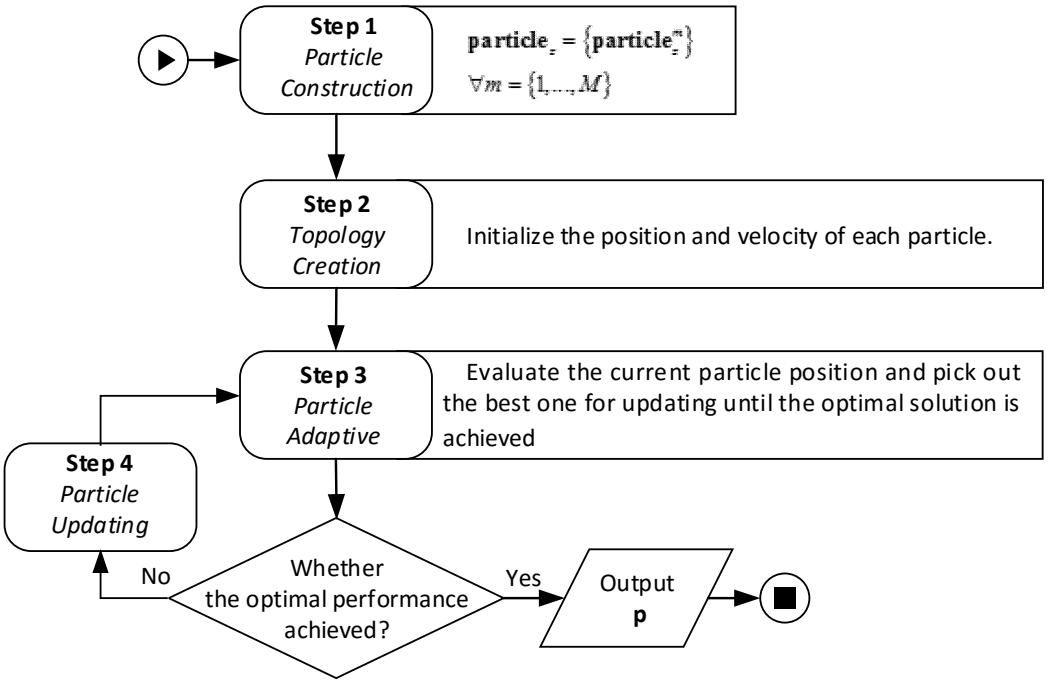


Fig. 3-6 Flowchart of PSO-PA

Step 1: (Particle Construction).

Assume there are Z^{PSO} particles. We denote vector

$$\mathbf{particle}_z=\{\mathbf{particle}_z^m\},\ m=\{1,\ldots,\ M\} \tag{3-28}$$

as the zth particle, where $\mathbf{particle}_z^m=\{\mathbf{particle}_z^m(n)\}$, $n=1,\ldots,\ N_m$ is an N_m dimension sub-vector that represents the power allocation in CC m, and each element $\mathbf{particle}_z^m(n)$ in the vector is a power value. The particular particle structure is illustrated in Fig. 3-7. Accordingly, we denote the vector $\mathbf{velocity}_z$ with the same structure as the velocity of the zth particle, and initialize it as the following expression:

$$\mathbf{velocity}_z^m(n)=p^{m,-}\left(\omega\text{-}0.5\right),\ \forall m,\ n,\ z \tag{3-29}$$

where ω follows standard uniform distribution in [0, 1].

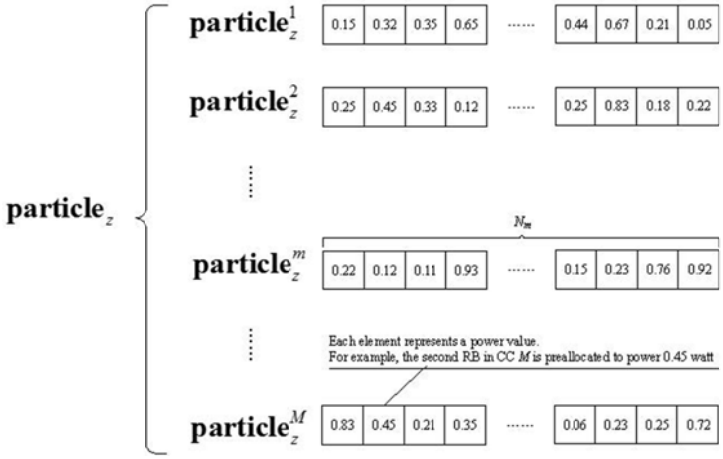


Figure 3-7 illustration of particle structure in PSO

STEP 2: (Topology Creation).

The target of topology creation is to initialize the position particles (the value of each element in $\mathbf{particle}_z$) and its velocity. We randomize all particles around equation power distribution by the following expression:

$$\mathbf{particle}_z^m(n)=p^{m,-}+p^{m,-}\left(\omega\text{-}0.5\right),\ \forall m,\ n,\ z \tag{3-30}$$

where ω follows standard uniform distribution in [0, 1].

STEP 3: (Particle Adapting).

The process of particle adapting is to evaluate the current particle position and pick out the best one for particle updating and the final solution. Each particle is comparable based on a fitness value that is related to the objective function and the constrain conditions. As to the problem ($\mathcal{P}3$) we define the fitness function of each

particle using the penalty function skill [65]:

$$F_z = \min_k R_k\left(\psi, \rho, \mathbf{particle}_z\right) \times \lambda_z \tag{3-31}$$

The fitness function is composed by two parts, where $\min_k R_k\left(\psi, \rho, \mathbf{particle}_z\right)$ indicates the minimal capacity of all UEs according to ψ, ρ and the current position of particle z, and λ_z is a penalty factor to ensure the constraint C3 can be satisfied. The expression of λ_z is

$$\lambda_z = 1 - \sqrt{\sum_{m=1}^{M}\left(\sum_{n=1}^{N_m} \mathbf{particle}_z^m(n) - P_{\text{total}}^m\right)^2} \tag{3-32}$$

Besides, we denote F^{best} and $\mathbf{Gparticle}^{\text{best}}$ as the known highest fitness value and the corresponding position, F_z^{best} and $\mathbf{Pparticle}^{\text{best}}$ as the known highest fitness value of particle *z* and the corresponding position, respectively.

STEP 4: (Particle Updating).

After adapting, we should update the particle position and its velocity to approach the excepted optimal solution. According to the suggestion in [66], the particles are manipulated according to the following equations:

$$\begin{aligned} \mathbf{velocity}_z^m &= \chi\Big(\mathbf{velocity}_z^m + c_1\omega_1\left(\mathbf{Pparticle}_z^{\mathbf{best}} - \mathbf{particle}_z^m\right) \\ &\quad + c_2\omega_2\left(\mathbf{Gparticle}^{\mathbf{best}} - \mathbf{particle}_z^m\right)\Big) \\ \mathbf{particle}_z^m &= \mathbf{particle}_z^m + \mathbf{velocity}_z^m \end{aligned} \tag{3-33}$$

where χ is inertia coefficient in PSO algorithm and the variables ω_1 and ω_2 are random variables with uniform distribution in interval [0, 1]. The relation between χ, c_1 and c_2 can be given as

$$\chi = \frac{2}{\left|2 - \left(c_1 + c_2\right) - \sqrt{\left(c_1 + c_2\right)^2 - 4\left(c_1 + c_2\right)}\right|} \tag{3-34}$$

where $c_1 = c_2 = 2.05$ and $\chi \approx 0.729$. The algorithm will repeat step (3) and (4) until the stopping criterion is reached. In this paper, we use the maximal iteration times T_{PSO} as the stopping criterion. The implementation of power allocation algorithm is summarized in **Algorithm 3-5**.

Algorithm 3-5 PSO based Power Allocation Algorithm

Input: K, M, N_m, $h_k^{m,n}$, Γ_k, γ_k, $\forall m,n,k$;

CC-allocation results ψ and RB-allocation results ρ;

Output: The power allocation results $\boldsymbol{p}$;

1: **for** $z \leftarrow 1$ to Z^{PSO} **do**

2: Initialize $\mathbf{particle}_z$ and $\mathbf{velocity}_z$ according to Eq. (3-29) and Eq. (3-30), respectively.

3: **end for**

4: **for** $t \leftarrow 1$ to T^{PSO} **do**

5: **for** $z \leftarrow 1$ to Z^{PSO} **do**

6: Calculate F_z according to Eq. (3-31).

7: **if** $F_z^{\mathrm{best}} < F_z$ **then**

8: $F_z^{\mathrm{best}} = F_z$, $\mathbf{Pparticle}_z^{\mathrm{best}} = \mathbf{particle}_{z^*}$

9: **end if**

10: **end for**

11: Find z^* that satisfies $F_{z^*} \geq F_z$, $\forall z$.

12: **if** $F^{\mathrm{best}} < F_{z^*}$ **then**

13: $F^{\mathrm{best}} = F_{z^*}$, $\mathbf{Gparticle}^{\mathrm{best}} = \mathbf{particle}_{z^*}$

14: **end if**

15: **for** $z \leftarrow 1$ to Z^{PSO} **do**

16: Updating $\mathbf{velocity}_z$ and $\mathbf{particle}_z$ according to Eq. (3-33).

17: **end for**

18: **end for**

3.5 Simulation

To evaluate the performance of the proposed algorithms, extensive simulations are carried out. The simulation layout is shown in Fig. 3-8. For large-scale path loss, the COST-231 Hata model is considered [67].

$$\begin{aligned} PL_{\mathrm{dB}} = &\left(44.9 - 6.55\log\left(h_{\mathrm{BS}}\right)\right)\log\left(d/1000\right) + 46.3 + \left(35.46 - 1.1h_{\mathrm{MS}}\right)\log\left(F_m\right) \\ &-13.82\log\left(h_{\mathrm{BS}}\right) + 0.7h_{\mathrm{MS}} + C \end{aligned} \tag{3-35}$$

where h_{BS} = 32m and h_{MS} = 1.5m are the height of eNB and UE respectively. The coverage radius of eNB is 1000m. There are 4 CCs in LTE-A system and the central frequency F_m equals to 700MHz, 1900MHz, 2300MHz and 3400MHz with bandwidth 20MHz, 10MHz, 10MHz, and 20MHz, respectively; the shadowing is implemented by lognormal distribution with standard deviation values of 6dB. For small-scale fading,

the Clarke's flat fading model with six independent Rayleigh multipath is adopted, the same as in [59]. The power delay profile is assumed exponentially decaying with e^{-2l}, where l is the multipath index. Hence, the relative power of the six multipath components is [0, −0.869, −17.37, −26.06, −34.74, −43.43]. Other parameters related to the following simulations are list in Table 3-2.

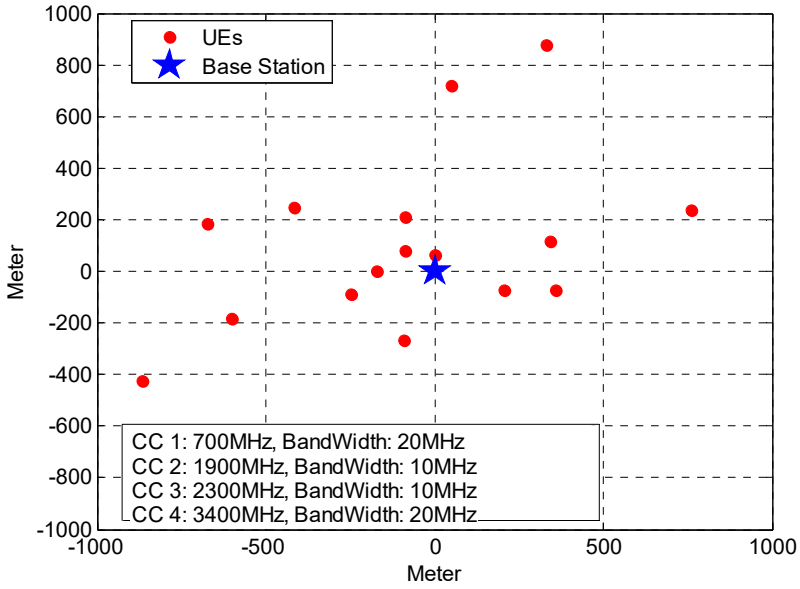


Figure 3-8 The simulation layout

Table 3-2 Simulation parameter setting

Parameter	Value
Maximal transmission power in CC 1 P^1_{total}	5Watt
Maximal transmission power in CC 2 P^2_{total}	10Watt
Maximal transmission power in CC 3 P^3_{total}	10Watt
Maximal transmission power in CC 4 P^4_{total}	10Watt
Power spectral density N_0	s10^{-17}
Bandwidth of each RB W	180kHz
Number of UEs K	16~32
Number of RFCs for each UE S	2
BER requirement BER	10^{-3}
Number of samples in PSO Z^{PSO}	100
Maximal number of iterations in PSO T^{PSO}	500

A. RB Allocation Estimation

To verify the effectiveness of the proposed RB allocation estimation approach, we compare actual allocation results with estimated allocation results by exploiting

proposed approach in this chapter in terms of sum capacity. As shown in Fig. 3-9, we can see that there is a little gap between actual allocation results and the approximate RB allocation approach, which indicates that the proposed RB allocation estimation approach is good enough to evaluate the CC allocation performance. In addition, it is observed that the proposed RB estimation approach can get a coarse fairness among UEs.

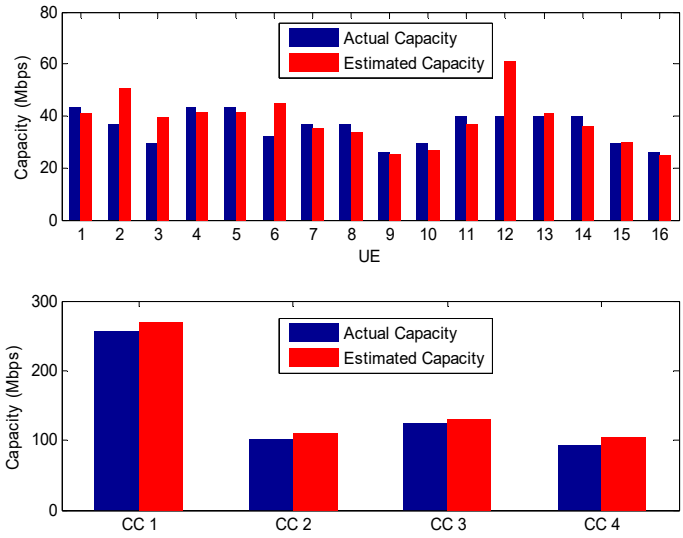


Figure 3-9 The effectiveness of proposed approximate RB allocation approach

B. CC Allocation

In Fig. 3-10, we present the convergence of proposed CEGA-CCA in terms of minimal UE's capacity, user fairness and the sum capacity. The start points represent the system performance by using GA-CCA algorithm. It can be seen that the GA-CCA algorithm can approach a relatively fair solution with $F > 0.95$. As the increase of iterations, the minimal UE's capacity (the objective of P1) of CC-allocation is improved by about 7.6% (from 33.77 Mbps to 36.51 Mbps). Besides, it is observed that there is a tradeoff relationship between UE's sum capacity and user fairness index, which means one increases usually as the other one decreases. This result has also been observed in other resources allocation related works.

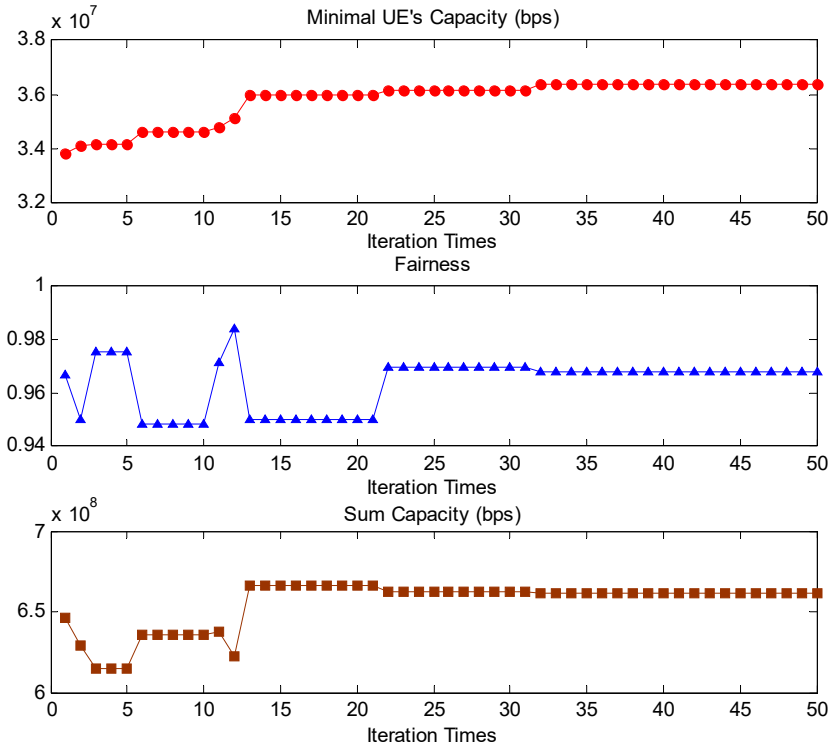


Figure 3-10 The convergence of proposed GA-CCA algorithm with different performance metrics

Table 3-3 Performance comparisons of different CC allocation algorithms

Algorithms	**Minimal UE's capacity**	**Fairness**	**Sum capacity**
Mobile Hashing	25.60Mbps	0.9186	604.49 Mbps
Round Robin	32.78Mbps	0.9252	614.53 Mbps
GA-CCA	33.77Mbps	0.9668	646.09 Mbps
CEGA-CCA	36.35Mbps	0.9679	661.45 Mbps

To further illustrate the efficiency and the effectiveness of the proposed CEGA-CCA algorithm, we compare the proposed GA-CCA and CEGA-CCA algorithms with the MH and the RR algorithms, as shown in Table 3-3. Due to the contribution of the proposed estimated RB allocation and equal power distribution in this chapter, the Mobile Hashing can also get a coarse fairness even through there is no extra balance effort. According to Table 3-3, we can see that the proposed GA-CCA and CEGA-CCA outperform Mobile Hashing and Round Robin in all aspects. Specifically, in contrast to the Mobile Hashing scheme, the proposed CEGA-CCA algorithm can get 40.4%, 5.4% and 9.4% improvement in terms of minimal UE's capacity, fairness and sum capacity, respectively.

C. RB Allocation

On the basis of the obtained CC-allocation result, we show the convergence of the proposed RE-RBA algorithm in Fig. 3-11, where the start points is the performance of the MG-RBA algorithm. As shown in Fig. 3-11, the RE-RBA converges to a stable solution with only several iterations. Furthermore, it is observed that all performance metrics can be improved with the increase of iterations, which indicates the capability of the proposed algorithm to drag the search out of the local extreme.

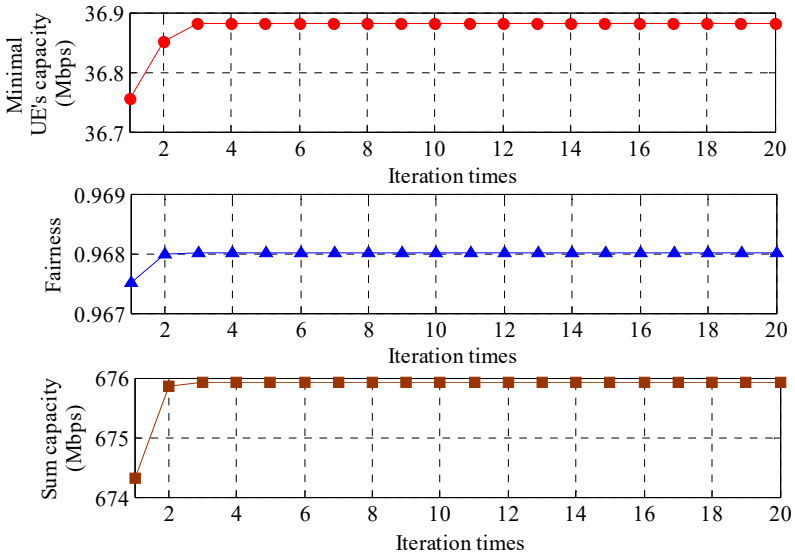


Fig. 3-11 Convergence of proposed RE-RBA algorithm with different performance metrics

Table 3-4 Utility function with optimal channel allocation

Algorithms	Minimal UE's capacity	Fairness	Sum capacity
CEGA-CCA	36.35 Mbps	0.9679	661.45 Mbps
CEGA-CCA + GA-RBA	36.71 Mbps	0.9675	674.32 Mbps
CEGA-CCA + RE-RBA	36.96 Mbps	0.9680	675.93 Mbps
CEGA-CCA+PSO-PA	37.38 Mbps	0.9782	643.09 Mbps
CEGA-CCA + GA-RBA + PSO-PA	37.96 Mbps	0.9769	655.62 Mbps
CEGA-CCA + RE-RBA + PSO-PA	38.08 Mbps	0.9777	657.47 Mbps

We also present the comparison results of CEGA-CCA without RB allocation, the GA-BRA and RE-BRA in Table 3-4. In contrast to the result without RB allocation, the proposed GA-RBA algorithm has better performance in terms of minimal UE's capacity and sum capacity, but slightly worse in terms of fairness. This is because the GA-RBA facilitates better use of multi-user diversity. As the enhanced method, the

proposed RE-RBA is superior to the others in all aspects. Compared to the non-RBA case, the RE-RBA algorithm can obtain 1.7%, 0.01% and 2.2% improvement in terms of minimal UEs capacity, fairness and sum capacity, respectively.

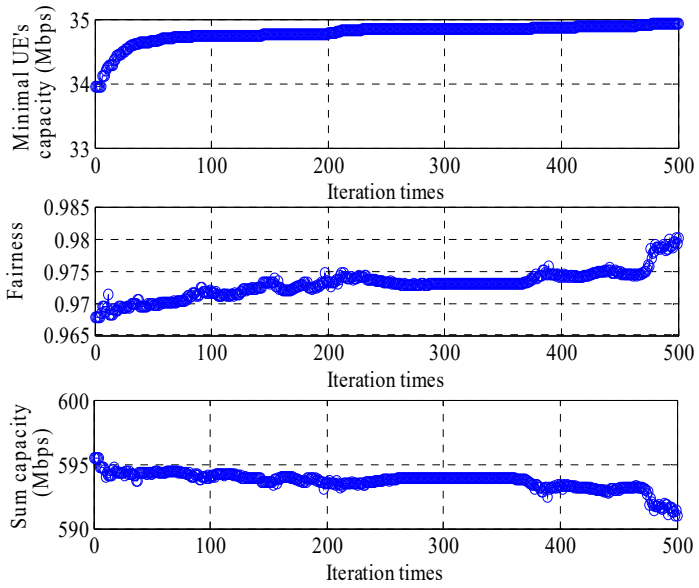


Fig. 3-12 The convergence of proposed PSO-PA algorithm with different performance metrics

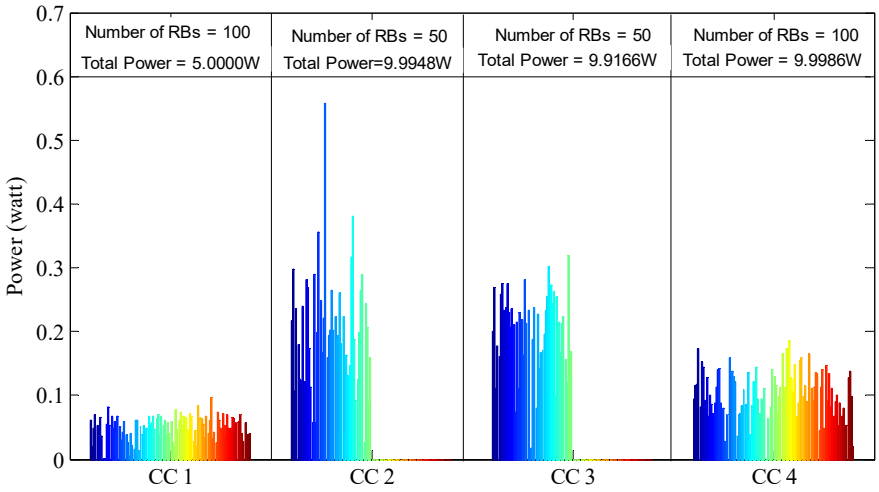


Figure 3-13 Power distributions after using the PSO-PA algorithm

D. Power allocation

Given the results of CC allocation and RB allocation, we can further improve the objective of problem ($\mathcal{P}3$) using the proposed PSO-PA algorithm. The performance of the proposed PSO-PA algorithm is shown in Fig. 3-12. As the increase of iterations, there is a slightly improvement in terms of minimal UE's capacity and fairness, but the

sum capacity decreases. This indicates that the power allocation just plays a fine adjustment role in all resource allocation process. We also show the detail power distribution in all RBs. As shown in Fig. 3-13, it can be seen that the power limitations in all CCs can be satisfied well, which verifies the effectiveness of our designed fitness function. As shown in Table 3-4, it is obvious that the system performance in terms of minimal UE's capacity and fairness is improved by the proposed PSO-PA algorithm.

It is worth noting that although all allocation stages can benefit the system performance, the levels of their contributions to the objective improvement are different. In CC-allocation stage, the minimal UE's capacity can be improve up to 40%, while this value is only 2% in both RB-allocation and power-allocation stages.

E. Complexity

In Fig. 3-14, we also present the time complexity of the proposed algorithms with the variation of *K*. It is seen that the time consumption of the three stages will increase linearly when the number of UEs increases, which indicates that the proposed algorithms are processable in polynomial time with respect to *K*.

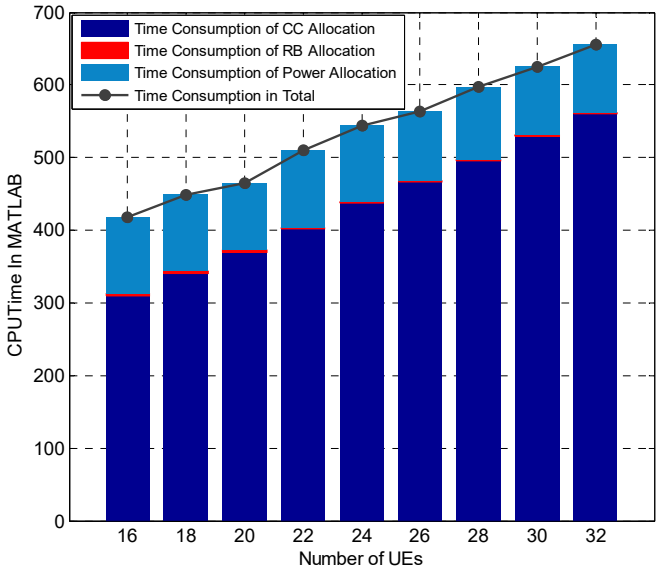


Figure 3-14 The time complexity of the proposed algorithms

3.6 Conclusion

CA is a promising technique to support higher data rate transmissions for the next generation wireless communication systems, while how to perform resource allocation with CA technique is still an open issue. In this paper, we have decomposed

the problem of resource allocation with CA into three subproblems, CC Allocation, RB Allocation, and power Allocation. For each subproblem, we have proposed low-complexity algorithms correspondingly. From the results, we have found that the proposed CC Allocation algorithm can get higher gains in terms of both minimal UE's sum capacity and proportional fairness, and the RB Allocation and the Power Allocation algorithms can also benefit to balance the tradeoff between the sum capacity and proportional fairness. However, we have also found that the CC allocation occupies most of the time consumption from our simulation results, and there is space remaining to lower the complexity of the CC allocation in future study.

Chapter 4 Energy-efficient radio resource management in HetNets with Coordinated Multiple Points

Energy-efficiency and spectrum-efficiency are the most important issues for future mobile systems. HetNets with CoMP are wildly approved as a promising solution to meet increasing demands of mobile data traffic and reduce energy consumptions. However, hyper-dense deployments and complex coordination mechanism introduce several challenges in RRM of mobile systems. To address this issue, we present an RRM approach for CoMP-based HetNets, which aims to maximize weighted energy efficiency while guaranteeing the data rate of each transmission. The proposed RRM approach is based on a CE optimization method that is an effective and low-complex heuristic algorithm. Furthermore, we also give the implementations of the proposed RRM approach in centralized and decentralized mode, respectively. At last, Extensive simulations are conducted to validate the effectiveness of the proposed schemes.

4.1 Research on energy-efficient resource allocation

Electric energy is one of the crucial elements that support the continuingly growing wireless mobile traffic. Energy consumption originating in wireless mobile communications increases the operating cost of vendors. Meanwhile, waste and environmental pollution caused by wireless mobile systems attracts wide attentions. The investigation report published by *Ofcom* shows that the energy consumption caused by the global Information and Communication Technology (ICT) industry accounts for about 3% of the all, i.e., around 600 TWh (1 TWh = 10^9 kWh), and the number will grow up to 1700 TWh by 2030 [68]. At the same time, energy consumption increases the CO2 remission, which exacerbates environmental damage. Therefore deeply researching on energy-efficient wireless communication system is not only the social responsibility of the ICT industry, but also a demand for its own development [69].

In 2009, Mobile Virtual Centre of Excellence (MVCE), an international research organization of UK, proposed an idea of Green Radio and kicked off the corresponding

research plan. In 2008, Celtic from Europe initiated a project named as Optimising Power Efficiency in mobile Radio Networks (OPERA-NET). OPERA-NET project contributed to improve the energy efficiency of wireless communication system at all aspects such as system, architecture and terminals, based on the high correlation between devices in current networks. In 2010, Energy Aware Radio and Network Technologies (EARTH) project [70][71], which is similar to Green Radio is, is started by a European organization FP 7 IP. The project is aiming at reducing energy waste in wireless wideband network without any loss of QoS. EARTH is one of the most significant research projects on energy-efficiency wireless communications systems. It has conducted the fundamental works for standardization development, efficiency evaluation, network architecture, performance optimization, prototype design and so on, which strongly influences future works [72]. In the same year, an alliance of several major mobile network operators, mobile network device vendors and research institutions from all over the world firmly launched Green Touch [70][71] project. According to the research outcome published in 2013, Green Touch project claimed that network energy consumption was able to fulfill a 90% reduction, by jointly optimizing network technologies, architecture, algorithms and protocols [73]. Addition to the projects mentions above, there are many other projects has been launched to deeply research on environment-friendly communication and computation techniques, as well as other relative fields, such as Communication Green (ComGreen) project from German, Green-T project started by European Celtic, Technical Committee on Green Communications & Computing (TCGCC) organized by IEEE, and so on.

4.2 Evaluation index sign system of energy efficiency

In mobile communication systems, energy efficiency reflects the tradeoff between system performance and required energy consumption. There are many different definitions of energy efficiency can be utilized in various scenarios. According to EARTH, energy efficiency in wireless communication systems can be categorized into component level, node level and network level [74].

Network components are the major contributors to energy consumption of BSs. Around the components of BSs, transceiver systems and baseband process units account for the main part. Fortunately, it is possible to adjust transceiver systems, as

well as baseband process units, to reduce energy consumption. This is one of the most attractive concerns in the related research area. Moreover, it is efficacious to reduce the total energy consumption by optimizing energy supplying systems and cooling systems of BSs.

Node-level energy efficiency is a ratio of effective energy output and total energy consumption of a BS and its accessories, such as backhaul links, Remote Radio Heads (RRH) and so forth. EARTH expresses the different meanings of node-level energy efficiency defined by Alliance for Telecommunications Industry Solutions (ATIS) and European Telecommunications Standards Institute (ETSI), respectively. The definition and method of ATIS are more general, while the definition of ETSI is more explicit. ETSI classifies power consumption of a BS according to the deployment mode (centralized or distributed), with consideration of traffic situations in busy time, middle time and idle time. As to a centralized BS, the power efficiency is defined as:

$$P_{\text{equipment}} = \frac{P_{\text{BH}} t_{\text{BH}} + P_{\text{med}} t_{\text{med}} + P_{\text{low}} t_{\text{low}}}{t_{\text{BH}} + t_{\text{med}} + t_{\text{low}}} \tag{4-1}$$

where P_{BH}, P_{med} and P_{low} are average power required in busy time, middle time and idle time, respectively; an t_{BH}, t_{med} and t_{low} are the corresponding time durations.

A similar definition of power efficiency in distributed system is

$$\begin{aligned} P_{\text{equipment}} &= P_{\text{C}} + P_{\text{RRH}} \\ &= \frac{P_{\text{BH,C}} t_{\text{BH}} + P_{\text{med,C}} t_{\text{med}} + P_{\text{low,C}} t_{\text{low}}}{t_{\text{BH}} + t_{\text{med}} + t_{\text{low}}} + \frac{P_{\text{BH,RRH}} t_{\text{BH}} + P_{\text{med,RRH}} t_{\text{med}} + P_{\text{low,RRH}} t_{\text{low}}}{t_{\text{BH}} + t_{\text{med}} + t_{\text{low}}} \end{aligned} \tag{4-2}$$

where P_{C} and P_{RRH} represent the power consumption caused by the control unit and RRH of a BS, respectively.

ETSI also provides a definition of side-level power efficiency, which integrating power consumption of the energy supplying system, as well as the cooling system, with node-level power efficiency by introducing a scale factor. Consequently, the centralized site-level power efficiency can be expressed as

$$P_{\text{site}} = PSF \cdot CF \cdot P_{\text{equipment}} \tag{4-3}$$

where *PSF* and *CF* are scale factors of the energy supplying system and the cooling system respectively. *PSF* and *CF* are constants without units in specific application scenarios.

Different from the centralized system, it is necessary to consider the loss due to energy supplying for RRH in a distributed system. Therefore, ETSI introduces a Power

Feeding Factor (PFF) into site-level power consumption expression in distributed systems

$$P_{\text{site}} = PSF \cdot CF \cdot P_{\text{equipment}} + PSF \cdot CF \cdot PFF \cdot P_{\text{RRH}} \tag{4-4}$$

According to the definitions provided by ETSI, ERTH proposes a method for quantifying energy efficiency (aslo known as Energy Consumption Index, shorted by ECI, in the reports.)

$$ECI = \frac{P_{\text{site}}}{\text{KPI}} \tag{4-5}$$

where P_{site} represents the power consumed by a site; KPI is an index that relates with the QoS level of users, and coverage and data rate of a BS is the most common uses as KPI in specific optimization problems.

To demonstrate more details, ECI can be further decomposed as

$$ECI = \frac{P_{\text{site}}}{\text{KPI}} = \frac{P_{\text{site}}}{P_{\text{equipment}}} \cdot \frac{P_{\text{equipment}}}{\text{KPI}} = \frac{1}{\eta_A} \cdot \frac{P_{\text{equipment}}}{\text{KPI}} = \frac{1}{\eta_A} \cdot \frac{P_{\text{equipment}}}{P_{\text{RF}}} \frac{P_{\text{RF}}}{\text{KPI}} = \frac{1}{\eta_A} \cdot \frac{1}{\eta_B} \frac{P_{\text{RF}}}{\text{KPI}} \tag{4-6}$$

where $1/\eta_A$ is the energy transformation ratio of the site; $1/\eta_B$ is the energy transformation ratio of the RRH; P_{RF}/KPI reflects the relationship between power of Radio Frequency (RF) and users' QoS.

Energy transformation in the site is illustrated in Fig. (4-1).

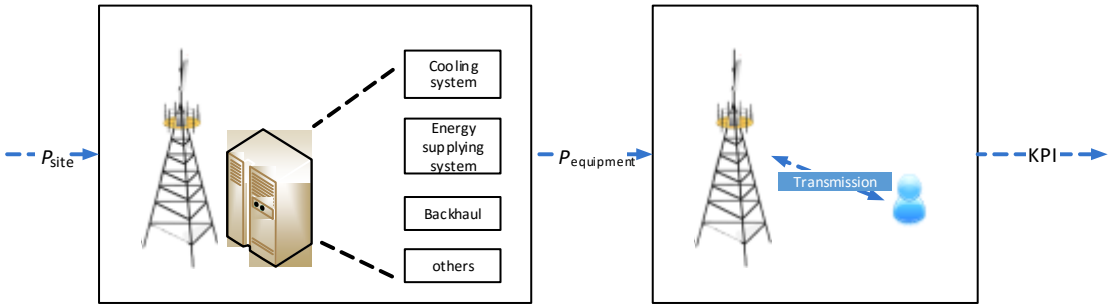


Fig. 4-1 Illustration of energy transformation

According to the decomposed expression of power consumption, we can separate the overall energy-efficiency improving problem into several subproblems, and independently optimize the system performance via $1/\eta_A$, $1/\eta_B$ and KPI. Usually, KPI is related to not only P_{RF}, but also the resource scheduling algorithm, power control algorithm, interference coordination methods, fairness requirement, capacity limit of backhaul links and so forth.

Eq. (4-6) gives a precise expression of energy efficiency at each node. However, we still cannot calculate the network-level energy efficiency by multiplying node-level energy efficiency by the number of nodes in the network, due to problems such as

intercell interference, handover, network traffic and wireless channel conditions. Moreover, it is not rational to consider network-level energy efficiency without consideration of network service qualities. Consequently, different definitions of energy efficiency are used at the network level. The widely used network-level energy efficiency includes: energy per bit information, power per unit of coverage and power per user.

1. Energy per bit information

Energy per bit information is defined as the ratio between energy consumption in the network and the amount of effective information it produced,

$$\lambda_1 = \frac{E}{I}\left(\text{J/bit}\right) \quad \text{or} \quad \lambda_1 = \frac{P}{R}\left(\text{Watt/bps}\right) \tag{4-7}$$

where E is the energy consumption of the network; I is the effective information transmitted under the consumed energy; P is the power consumption of the network; and R is the obtained data rate.

Energy per bit information is the most common evaluation index sign of wireless communication systems in the academic field. It also be used for theoretical analysis and energy efficiency evaluation for a single link, since both energy consumption of network and the amount of effective information is capable to be explicitly output. The major defect of energy per bit information is the lack of considerations like the mutual influence between neighboring cells, fairness and so forth. As the result, this evaluation index sign could not comprehensively reflect the energy efficiency of the network.

2. Power per unit of coverage

Power per unit of coverage is a ratio of power consumption and the coverage area, which can be expressed as

$$\lambda_A = \frac{P}{A}\left(\text{Watt/m}^2\right) \tag{4-8}$$

where A is the area of coverage in the unit of m^2.

The amount of CO_2 emission can be easily calculated according to power per unit of coverage. In addition, this evaluation index can also be used as one of the parameters accessing coverage performance of cellular networks. It is noticeable that there is an implicit relationship between network coverage and effective transmission of information. The coverage of a network can usually be determined by SINR of the user access and the sensitivity of user receivers. This means that SINRs within the coverage must be higher than a given threshold. Since higher SINR leads to better

data rate, the amount of information transmitted effectively is large with higher SINRs.

3. Power per user

Power per user is the ratio of power consumption and the number of users being served,

$$\lambda_{\mathrm{A}} = \frac{P}{N_s}\left(\mathrm{Watt/user}\right) \tag{4-9}$$

This is an intuitive evaluation index for mobile network operators. It is also suitable to be used to evaluation Quality of Service (QoS), Quality of Experience (QoE) and other key metrics of the network. However, the diversity of user types is ignored in this evaluation index, which makes it is not very useful for theoretical analysis.

Beside the evaluation index mentioned above, compositions that integrating the three network-level energy efficiency are also proposed, such as power /cell/unit area/user.

4.3 System Model

For the sake of simplicity, each independent transmitter, including the macro cell site and micro cell sites, is referred to as a Transmit Point (TP) in this paper. We consider a downlink system in a CoMP-based HetNet with M TPs and K user elements (UEs), as shown in Fig. 4-2. N_{T} and N_{R} represent the number of antennas on each TP and each UE, respectively. A control unit (CU) is assumed to be located at the center of the network, which is responsible for managing data information, as well as collecting all the CSI in this network. TPs are connected with the CU by backhaul links, through which control information and data packets are delivered to TPs from CU. Since JT CoMP is employed in our work, there are a large number of data packets need to be transmitted via backhaul links.

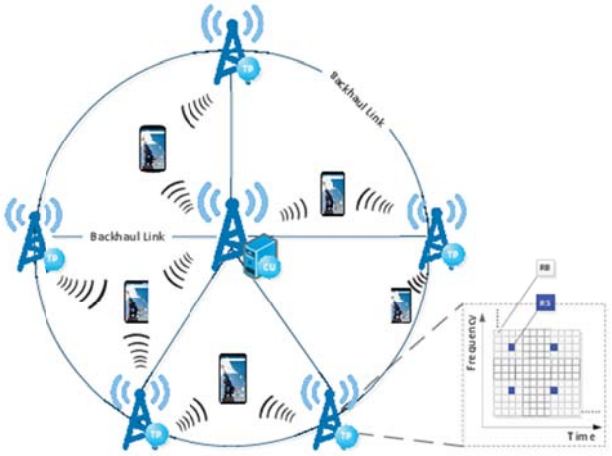


Figure 4-2 System model

The unit of radio resource in both time and frequency dimensions is referred to as RBs. As defined in LTE standard, an RB consists of 12 consecutive subcarriers for a duration of a transmit time interval (TTI) [75]. In this chapter, we assume all TPs share the same spectrum bandwidth, which is divided into totally N_{RB} RBs. Important notations used in this chapter are listed in Table 4-1. Several other important assumptions are considered in our work: Channel fading is considered to be quasi-static, so that channel coefficients remain constant during per TTI; perfect CSI acknowledgement is assumed at both receivers and transmitters; TPs are synchronized in terms of time, frequency and phase, which is reasonable in the considered system thanks to backhaul connections.

Table 4-1 Important denotations

Denotations	Meaning
N / K	number of TPs / UEs
N_T / N_R	number of antennas on each TP / UE
$\mathcal{M}$	the set consists of all the TPs
$\mathcal{M}_k$	CoMP set of UE k
$\beta_{m,k}^n$	scheduling index
$\mathcal{U}_m$	the set consists of UEs that attaching to TP m
R_k^n	data rate of UE k on RB n
$\overline{R}_k$	accumulated average data rate of UE k
p_m^n	transmit power used at TP m on RB n
N_{RB}	number of RBs

4.3.1 CoMP set selection

Ideally, a UE can achieve the optimal data rate if all TPs cooperatively transmit to it. However, the corresponding power consumption and computational complexity is unaffordable. An alternative is to select a CoMP set for the UE according to channel conditions. TPs in the CoMP set can provide the UE a favorable data rate at a much lower cost. A UE-specific selection of CoMP set includes three steps:

1. TPs deliver reference signals (RSs) periodically.
2. A UE hears the channels and measures them according to the strength of the received RSs. Based on the measurements and a given selection rule, the UE decides its own CoMP set.
3. The UE acknowledges its decision to the CU.

Denote $\mathcal{M}_k$ as the CoMP set of UE *k*, and $\mathcal{M}$ as the set including all TPs in the network. UE *k* decides its $\mathcal{M}_k$ following the rule below:

$$\begin{cases} m \in \mathcal{M}_k & m = \arg_{\mathcal{M}} RS(\text{the strongest}) \quad \text{or} \\ & RS(\text{the strongest}) - RS(\text{TP} \quad m) \le \Delta \\ m \notin \mathcal{M}_k & \text{otherwise} \end{cases} \tag{4-10}$$

where *RS* indicates the strength of the reference signal, and Δ is a threshold in dB. UE *k* distinguishes the strongest RS at the first place, and adds the corresponding TP to the CoMP set $\mathcal{M}_k$. Other TPs will be added to $\mathcal{M}_k$ only if strength of their RSs is no less than RS(the strongest) -Δ dB. As suggested in LTE releases, a rational Δ is in the range of 5~6 dB [76].

In the case where $\mathcal{M}_k$ includes only 'the strongest TP', UE *k* is referred to as a non-CoMP UE, since no cooperation occurs during downlink transmissions towards it. On the contrary, UE *k*' is referred to as a CoMP UE if its CoMP set $\mathcal{M}_k$ includes more than one TPs. A CoMP UE is possibly is located at an overlapping area of neighboring cells where intercell interference seriously damages transmissions. To combat with interference, TPs in $\mathcal{M}_{k'}$ are asked to conduct JT CoMP transmissions to UE *k*' for strengthening transmit signals and reducing intercell interference.

4.3.2 Dynamic JT CoMP transmission

JT CoMP allows TPs in UE's CoMP set to simultaneously transmit the desired data signal to it. Due to spatially separation of transmit antennas, multiple versions of the

desired signal will be received by the UE, which generates extra spatial diversity gain and strengthens the signal. The obtained data rate of a JT CoMP transmission to UE *k* on RB *n* is given as

$$R_k^n = b\log\left(1+\frac{\sum_{m\in\mathcal{M}_k}\left\|\mathbf{H}_{m,k}^n\mathbf{w}_m^n\right\|^2 p_m^n}{\sum_{m\in\mathcal{M}\setminus\mathcal{M}_k}\left\|\mathbf{H}_{m',k}^n\mathbf{w}_{m'}^n\right\|^2 p_{m'}^n+\sigma^2}\right) \tag{4-11}$$

where $\mathbf{H}_{m,k}^n$ is an $N_R \times N_T$ channel matrix between TP *m* and UE *k* on the *n*th RB. $\mathcal{M}\setminus\mathcal{M}_k$ is the complementary of $\mathcal{M}_k$ in $\mathcal{M}$, which includes all the interfering TPs. $\mathbf{w}_m^n$ is the precoding vector with dimension of $N_T \times 1$, which maps data streams onto the transmit antennas of TP *m*, and $\left(\mathbf{w}_m^n\right)^{\mathrm{H}}\mathbf{w}_m^n = 1, \forall m,n$. p_m^n is the power used by TP *m* for transmitting on RB *n*, and $\mathbf{n}_k^n \sim \mathcal{CN}\left(0,\sigma^2\mathbf{I}_{N_R}\right)$ is the corresponding complex Gaussian noise vector. *b* represents the bandwidth of an RB, which is standardized to be 180 kHz in LTE-Advanced systems [75].

Eq. (4-11) implies a static coordinated strategy where TPs in $\mathcal{M}_k$ are all required to serve UE *k* all the time. However, static strategies are not always optimal due to the time variation of wireless channels. To further improve the network performance, we use a dynamic JT CoMP strategy where a subset of each $\mathcal{M}_k$ rather than $\mathcal{M}_k$ is adaptively determined to perform JT CoMP transmission. Define a scheduling index $\beta_{m,k}^n \in \{0,1\}$ to indicate scheduling results, where $\beta_{m,k}^n = 1$ means that TP *m* is chosen to transmit to UE *k* on the *n*th RB. Then the data rate of a dynamic JT transmission can be given as

$$R_k^n = b\log\left(1+\frac{\sum_{m=1}^{M}\beta_{m,k}^n\left\|\mathbf{H}_{m,k}^n\mathbf{w}_m^n\right\|^2 p_m^n}{\sum_{m'=1}^{M}\left(1-\beta_{m',k}^n\right)\left\|\mathbf{H}_{m',k}^n\mathbf{w}_{m'}^n\right\|^2 p_{m'}^n+\sigma^2}\right) \tag{4-12}$$

where $\sum_{m=1}^{M}\beta_{m,k}^n \geq 1$.

4.3.3 Problem Formulation

In this chapter, we consider a practical RRM problem in terms of both spectrum and

power in a CoMP-based HetNet modeled in the last section. To improve synthesis performance, the objective involves data rates, power consumption and fairness among UEs at the same time.

The optimal data rate of the network can be achieved if resources are allocated to UEs with better channel conditions, regardless of those in 'poor' condition. A side-effect of this scheme is unfavorable fairness of UEs' data rates. In order to improve the fairness, we introduce the concept of proportional fairness into the objective of the RRM problem. As in [77], we weighted UE's data rate by its average data rate defined by

$$\overline{R}_k = \alpha \overline{R}_k^{\text{before}} + (1-\alpha) R_k \tag{4-13}$$

where $0<\alpha<1$ is the forgetting factor. The introduced weights bring UEs with worse channel conditions more possibility to occupy resources, and therefore increase the UEs' data rates.

Additionally, for the purpose of conserving energy, energy efficiency should be thoughtfully considered. The energy efficiency is defined as the ratio of obtained data rate to the total power consumed correspondingly. Combining with proportional fairness principle, the objective of the RRM problem is formulated as,

$$\max \quad \sum_{k=1}^{K}\sum_{n=1}^{N_{\text{RB}}} \frac{R_k^n}{\overline{R}_k} \Bigg/ \sum_{m=1}^{M}\sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n p_m^n \tag{4-14}$$

In a practical network, system performance is restricted by several factors. In addition to limited transmit power at each TP, finite capacity of backhaul links defines the upper limit of throughput achieved by a TP during a TTI. Furthermore, to guarantee quality of transmissions, we impose a threshold to data rate of each transmission. In summary, the RRM problem of the considered system can be formulated as

$$\begin{aligned}
\max \quad & \sum_{k=1}^{K}\sum_{n=1}^{N_{\text{RB}}} \frac{R_k^n}{\overline{R}_k} \Bigg/ \sum_{m=1}^{M}\sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n p_m^n \\
\text{s.t.} \quad & \text{C1} \quad 0 \le p_m^n \le S, \forall m,n \\
& \text{C2} \quad \beta_{m,k}^n \in \{0,1\}, \sum_{k=1}^{K} \beta_{m,k}^n \le 1, \forall m,n,k \\
& \text{C3} \quad \sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n R_k^n \le C_m, \forall m \\
& \text{C4} \quad R_k^n \ge R_{\text{thres}}, \forall k,n
\end{aligned} \tag{4-15}$$

In Eq. (4-15), C1 shows the power constraint at each transmission where S is the

largest transmit power allowed by the system; C2 ensures each index $\beta_{m,k}^n$to be a bit number, so that a TP can serve no more than one UE on each RB; C3 demonstrates the constrained throughput of a TP caused by limited capacity of backhaul connections to the CU, where C_m represents the capacity of the backhaul link connecting TP m and the CU; and C4 guarantees the data rate of each ongoing transmission, where R_{thres} is the given threshold of data rate.

4.4 Centralized solution

Since the problem in Eq. (4-15) is NP-hard, it is unpractical to achieve the optimal solution. An alternative method is to consider the problem as a combination of a scheduling problem under the consumption of equal power allocation and a power allocation problem with given scheduling result. In this way, an approximated solution to the problem can be obtained in a polynomial time. In the rest of this section, we propose a heuristic algorithm based on CE for RB scheduling with equal transmit power at the first place. Then a KKT-method to solve the power allocation problem is presented.

The algorithm involving both RB scheduling and power allocation proposed in this section is centralized, which means that the resource allocation is operated at the CU with global CSIs. The centralized algorithm is capable to achieve a favorable system performance, but it requires significant computational effort of the CU.

4.4.1 CE-Based scheduling algorithm

First, we propose a RB scheduling algorithm based on CE method. The objective of the considered RB scheduling problem becomes

$$\max_{\beta_{m,k}^n, p_m^n} \quad \sum_{k=1}^{K}\sum_{n=1}^{N_{\text{RB}}} \frac{R_k^n}{\overline{R}_k} \Bigg/ \sum_{m=1}^{M}\sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n S \tag{4-16}$$

where S is the fixed transmit power. The objective is constraints by C2-C4 in (4-15). We first propose a CE-based algorithm to solve the RB scheduling problem described above.

CE method is a typical heuristic algorithm to estimate probabilities of rare events in complex stochastic networks [78], and to deal with linear programming. The basic idea of CE method is to generate sufficient samples under a given strategy, and then

update the generating strategy according to samples. After iteratively repeating this procedure, generated samples will converge to the optimal solution. The proposed CE-based scheduling algorithm follows three major stages including initialization, iteration and a complementary stage to close unfavorable transmissions, as illustrated in Fig. 4-3.

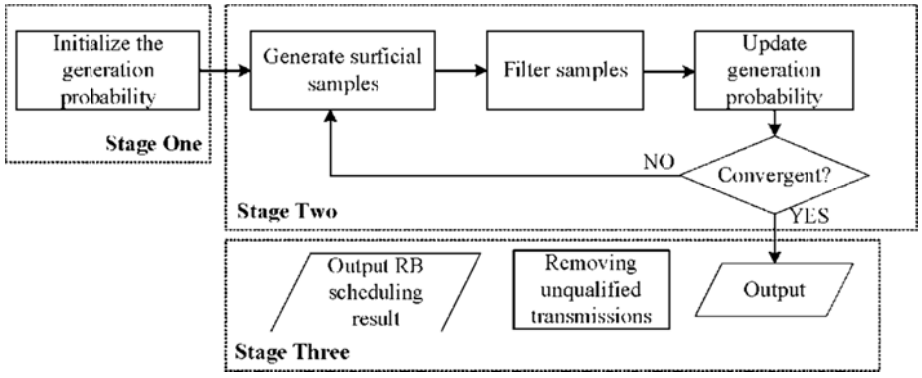


Figure 4-3 The flowchart of CE-based centralized RB scheduling algorithm

1 Initialization

LTE $\mathbf{X}^m=\left[x_{(1)}^m,\ldots,x_{(n)}^m,\ldots,x_{(N_{\mathrm{RB}})}^m\right]\left(x_{(n)}^m \in [0,\mathcal{U}_m]\right)$ denote a sample generated according to a given probability, which presents a possible scheduling result of TP m. $x_{(n)}^m = 0$ means that no transmission is scheduled on RB n of TP m, otherwise $x_{(n)}^m$ is the ID of the scheduled UE. In the CE method, the scheduling problem is regarded as a stochastic procedure. The distribution of $x_{(n)}^m$ is denoted by $\mathbf{q}^{m,n}=\left[q_0^{m,n},q_1^{m,n},\ldots,q_u^{m,n},\ldots,\mathrm{q}_{|\mathcal{U}_m|}^{m,n}\right]$, where $q_u^{m,n}=\mathbb{P}\{x_{(n)}^m=u\}$, $u \in [0,\mathcal{U}_m]$. Obviously, $q_u^{m,n}$ has the attributes of $0\le q_u^{m,n}\le 1, \forall u,m,n$ and $\sum_{u\in[0,\mathcal{U}_m]} q_u^{m,n} = 1, \forall m, n$.

For the sake of accelerating convergence, we design the initial probability distribution according to estimated data rates. We first defined $\mathrm{P}\{x_{(n)}^m=0\} = \mathrm{Pr}_0$ where Pr_0 is a given value. This probability gives an opportunity to TP m to scheduling no transmission on RB n. For each $u \in \mathcal{U}_m$ the data rate R_u^n is estimated based on the fixed JT CoMP transmission as given by (4-11) at the initialization stage. As suggested by C4 in (4-15), the transmission is considered to be unqualified if $R_u^n < R_{\mathrm{thres}}$. For the rest of optional UEs, the probability is defined as $q_u^{m,n} = (R_u^n/R_{\mathrm{tot}}^m) \times \left(1\text{- }\mathrm{Pr}_0\right)$ where R_{tot} indicates the sum data rate of all possible qualified transmissions, and defined as

$$R_{tot}^{m} = \sum_{k \in \mathcal{U}_m, R_k^n \geq R_{\text{thres}}} R_k^n \tag{4-17}$$

Summarily, the initial distribution of $x_{(n)}^{m}$ is

$$q_u^{m,n} = \begin{cases} \text{Pr}_0, & u = 0 \\ \left(1 - \text{Pr}_0\right) \times R_u^n / R_{\text{tot}}^m, & u \in \mathcal{U}_m, \quad \text{and} \quad R_u^n \geq R_{\text{thres}} \\ 0, & u \in \mathcal{U}_m, \quad \text{and} \quad R_u^n < R_{\text{thres}} \end{cases} \tag{4-18}$$

2 Iteration

Each iteration of the proposed algorithm includes three steps. First, the algorithm generates adequate samples according to a given strategy. Then it is necessary to exclude samples those are not satisfied the constraints, and to select 'good' samples for the next stage. At last, probability distribution needs to be updated according to the selected samples, so that 'better' samples will be generated in the next iteration. After sufficient iterations, the algorithm gradually approaches the optimal solution.

Let N_{SAM} denote the number of samples generated in each iteration for each TP, and $\mathbf{X}_1^m, \ldots, \mathbf{X}_i^m, \ldots, \mathbf{X}_{N_{\text{SAM}}}^m$ denote the corresponding samples, where $\mathbf{X}_i^m = \left[x_{i(1)}^m, \ldots, x_{i(n)}^m, \ldots, x_{i(N_{\text{RB}})}^m\right]$. Each sample can map into a scheduling index set $\left\{\beta_{m,k}^n, \forall k, n\right\}$, and leads to a weighted energy efficiency of TP m given as

$$f_m\left(\mathbf{X}_i^m\right) = \sum_{n=1}^{N_{\text{RB}}} \frac{R_{x_{i(n)}^m}^n}{\overline{R}_{x_{i(n)}^m}} \Bigg/ \sum_{n=1, x_{i(n)}^m \neq 0}^{N_{\text{RB}}} S \tag{4-19}$$

where $\sum_{n=1, x_{i(n)}^m \neq 0}^{N_{\text{RB}}} S$ represents the total power consumption at TP m according to sample $\mathbf{X}_i^m$. The sum data rate of TP m can be estimated as $R_m(\mathbf{X}_i^m) = \sum_{n=1}^{N_{\text{RB}}} R_{x_{i(n)}^m}^n$.

A qualified sample should satisfy two requirements. First, the sum data rate R_m cannot exceed the backhaul capacity of TP m. Therefore, samples those lead to overlarge sum data rates, i.e., $R_m(\mathbf{X}_i^m) > C_m$, will be removed. Second, the value of $f_m(\mathbf{X}_i^m)$ should be high enough. Samples whose weighted energy efficiency $f_m(\mathbf{X}_i^m) < f_{\text{thres}}^{(t)}$ will also be removed, where $f_{\text{thres}}^{(t)}$ is a threshold that increases after each iteration until it converges. Consequently, N_{IM} ($N_{\text{IM}} \leq N_{\text{SAM}}$) qualified samples are left for updating the generation probability. Without loss of generality, qualified samples are denoted as $\mathbf{X}_j^m (1 \leq j \leq N_{\text{IM}})$.

At the last stage of iteration, possibility distributions are updated on the basis of qualified samples in order to generate better samples in the next iteration. The

updated possibility $\mathbb{P}\{x_{(n)}^m=u\}$ is

$$q_u^{m,n}=N\left(x_{j(n)}^m=u\right)\Big/N_{\text{IM}} \tag{4-20}$$

where $N\left(x_{j(n)}^m=u\right)$ represents the number of times that UE u appears in N_{IM} samples on the nth RB.

The proposed iteration algorithm is summarized as the **Algorithm 4-1**.

Algorithm 4-1 Iteration in CE-based scheduling algorithm

1: $f_{\max}=f_{\max\_pre}=0$, $counter=1$
2: **while** $counter \le N_{\text{SAM}}$ **do**
3: Generate samples $\mathbf{X}^m$ according to the distribution $\mathbf{q}^{m,n}$.
4: Calculate utility function of $\mathbf{X}_i^m$, i.e., $f(\mathbf{X}_i^m)$, according to (3-9).
5: Calculate the sum data rate of $\mathbf{X}^m$, i.e., $R_m(\mathbf{X}^m)=\sum_{k\in\mathcal{U}_m}R_k^n$.
6: **if** $f(\mathbf{X}_i^m)<f_{\text{thres}}^{(t)}$ **then**
7: **CONTINUE;**
8: **end if**
9: **if** $R_m(\mathbf{X}^m)>C_m$ **then**
10: **CONTINUE**;
11: **end if**
12: $\mathbf{X}_{counter}^m=\mathbf{X}^m$
13: $counter=counter+1$
14: **end while**
15: **for** $i=1$ to N_{SAM} **do**
16: Calculate $f(\mathbf{X}_i^m)$ according to (3-9)
17: **end for**
18: Sort samples in a descending order in terms of $f(\mathbf{X}_i^m)$. Denote the consequence by $\hat{\mathbf{X}}_1^m,\hat{\mathbf{X}}_2^m,\ldots,\hat{\mathbf{X}}_{\text{SAM}}^m$.
19: Calculate $N_{\text{IM}}=\lceil(1-\rho)N_{\text{SAM}}\rceil$, and let $f_{\text{thres}}^{(t+1)}=f\left(\hat{\mathbf{X}}_{N_{\text{IM}}}^m\right)$
20: **if** $f\left(\hat{\mathbf{X}}_{N_{\text{IM}}}^m\right)>f_{\max}$ **then**
21: $\mathbf{X}_{\text{out}}^m=\hat{\mathbf{X}}_1^m$, $f_{\max\_pre}=f_{\max}$, $f_{\max}=f\left(\hat{\mathbf{X}}_1^m\right)$
22: **end if**

23:Update $\mathbf{q}^{m,n}$ according to (4-20)

24:Map $\mathbf{X}_{\text{out}}^{m}$ into $\beta_{m,k}^{n}$, and recalculate R_k^n according to (3-3).

3 Removing unqualified transmissions

The algorithm proposed above cannot ensure that data rate of each transmission is as high as the given threshold R_{thres}. Since we put our purpose on improve energy efficiency, it is reasonable to close transmissions those are estimated to be unqualified.

Algorithm 4-2 summaries the entire RB scheduling algorithm proposed above.

Algorithm 4-2 Centralized RB scheduling algorithm

1: Calculate R_k^n according to Eq. (4-11).
2: Initialized probability distribution $\mathbf{q}^m$ ($\forall m$) according to Eq. (4-18).
3: **for** t = 1:$t_{\max}$ **do**
4: **for** m = 1:M **do**
5: Process **Algorithm 1**, and output $\mathbf{X}_{\text{out}}^{m}(\forall m)$.
6: **end for**
7: Map obtained $\mathbf{X}_{\text{out}}^{m}$ into $\left\{\beta_{m,k}^{n}\right\}$.
8: Update R_k^n according to Eq. (3-3) and obtained $\left\{\beta_{m,k}^{n}\right\}$.
9: **if** (all the elements in $\mathbf{q}^m$ ($\forall m$) converge) **then**
10: Output $\mathbf{X}_{\text{out}}$
11: **BREAK**
12: **end if**
13:**end for**
14: Map obtained $\mathbf{X}_{\text{out}}^{m}$ into $\left\{\beta_{m,k}^{n}\right\}$
15:**for** n=1:N_{RB} **do**
16: **for** u=1:K **do**
17: **if** R_k^n<R_{thres} **then**
18: Let the corresponding $\beta_{m,u}^{n} = 0$
19: **end if**
20: **end for**
21:**end for**

4.4.2 Power allocation algorithm

The considered power allocation problem is non-convex, for which optimal solution does not exist. In this section, we use an analytical method on the base of KKT-condition to approach a local optimal solution of power allocation. To simply the problem, we intend to decompose the problem into independent subproblems with much less variables. The overall power allocation algorithm is illustrated in Fig. 4-4.

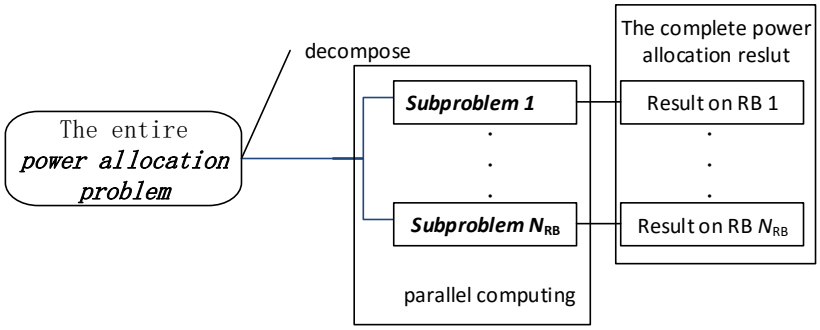


Figure 4-4 The flowchart of KKT-based power allocation algorithm

With the obtained RB scheduling result, the power allocation problem can be given as,

$$\begin{aligned} \max_{\beta_{m,k}^n, p_m^n} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N_{\mathrm{RB}}} \frac{R_k^n}{\overline{R}_k} \Big/ \sum_{m=1}^{M}\sum_{n=1}^{N_{\mathrm{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n p_m^n \\ \text{s.t.} \quad & \text{C1:} \quad 0 \le p_m^n \le S, \forall m,n \\ & \text{C3:} \quad \sum_{n=1}^{N_{\mathrm{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n R_k^n \le C_m, \forall m \\ & \text{C4:} \quad R_k^n \ge R_{\mathrm{thres}}, \forall k,n \end{aligned} \tag{4-21}$$

The objective function in Eq. (4-21) is a well-known non-convex function for which optimal solution does not exist. In this work, we use an analytical method on the base of KKT-condition to approach a local optimal solution of power allocation.

A solution of Eq. (4-21) can be achieved by solving its dual function in terms of backhaul and data rate constraints given as

$$\begin{aligned} \max_{p_m^n} \quad & \sum_{n=1}^{N_{\mathrm{RB}}} \left(\sum_{k=1}^{K} \frac{R_k^n}{\overline{R}_k} \Big/ \sum_{m=1}^{M} p_m^n \right) - \sum_{m=1}^{M} \lambda_m \left(\sum_{n=1}^{N_{\mathrm{RB}}}\sum_{k=1}^{K} \beta_{m,k}^n R_k^n - C_m \right) + \sum_{n=1}^{N_{\mathrm{RB}}}\sum_{k=1}^{K} \mu_{n,k} \left(R_k^n - R_{\mathrm{thres}} \right) \\ \text{s.t.} \quad & 0 \le p_m^n \le S, \forall m,n \end{aligned} \tag{4-22}$$

where $\{\lambda_m, \forall m\}$ and $\{\mu_{n,k}, \forall n,k\}$ are non-negative Lagrangian multipliers. However, the dual function given above is still hard to be solved since it involves too much variables (MN_{RB} variables and Lagrangian multipliers). To further simply the problem, we intend to decompose Eq. (4-22) into independent subproblems with fewer

variables.

Let $\beta_m = \sum_{n=1}^{N_{\mathrm{RB}}} \sum_{k=1}^{K} \beta_{m,k}^n$. $\beta_m = 0$ indicates the situation where non transmission is scheduled for TP *m*. This is unreasonable in the high-loading network we considered. Thus we suppose that $\beta_m \neq 0$, consequently C_m can be rewritten as

$$C_m = \sum_{n=1}^{N_{\mathrm{RB}}} \sum_{k=1}^{K} \beta_{m,k}^n \frac{C_m}{\sum_{n=1}^{N_{\mathrm{RB}}} \sum_{k=1}^{K} \beta_{m,k}^n} = \sum_{n=1}^{N_{\mathrm{RB}}} \sum_{k=1}^{K} \beta_{m,k}^n \frac{C_m}{\beta_m} \tag{4-23}$$

After substituting Eq. (4-23) into Eq. (4-22), the problem can be decomposed into N_{RB} independent subproblems, where each subproblem is given as

$$\begin{aligned} \max_{p_m^n} \quad & \sum_{k=1}^{K} \frac{R_k^n}{\overline{R}_k} \Big/ \sum_{m=1}^{M} p_m^n - \sum_{k=1}^{K} \sum_{m=1}^{M} \lambda_m \beta_{m,k}^n \left(R_k^n - \frac{C_m}{\beta_m} \right) + \sum_{n=1}^{N_{\mathrm{RB}}} \sum_{k=1}^{K} \mu_{n,k} \left(R_k^n - R_{\mathrm{thres}} \right) \\ \text{s.t.} \quad & 0 \le p_m^n \le S, \forall m \end{aligned} \tag{4-24}$$

Each subproblem in Eq. (4-24) involves *M* variables with Lagrangian multipliers only, and can be solved independently on each RB. The computational complexity is significantly reduced in this way. In the rest of this subsection, we proposed an iterative method to solve each subproblem.

Let $f_n\left(p_m^n, \lambda_m, \mu_{n,k}\right)$ denote the objective function of Eq. (4-24). Take the first order derivative in terms of p_m^n and make it equal to zero, then a possible value of power allocation, denoted by $\hat{p}_m^n$, can be obtained as follows,

$$\hat{p}_m^n = \left\{ \frac{\left(\frac{1}{\overline{R}_{k^*} p_{\mathrm{tot}}^n} - \sum_{m=1}^{M} \lambda_m \beta_{m,k^*}^n + \mu_{n,k^*} \right)}{\frac{1}{\left(p_{\mathrm{tot}}^n \right)^2} \sum_{k=1}^{K} \frac{R_k^n}{\overline{R}_k} - \sum_{k=1,k\neq k^*}^{K} \left(\frac{1}{\overline{R} p_{\mathrm{tot}}^n} - \sum_{m=1}^{M} \lambda_m \beta_{m,k}^n + \mu_{n,k} \right) \times \frac{\partial R_k^n}{\partial p_m^n}} - A_m^n \right\} \times \frac{1}{\left\| \mathbf{H}_{m,k^*}^n \mathbf{w}_m^n \right\|^2} \tag{4-25}$$

where $p_{\mathrm{tot}}^n = \sum_{m=1}^{M} p_m^n$, and k^* is the UE that is scheduled on RB *n* of TP *m*, i.e., $\beta_{m,k^*}^n = 1$. The derivative $\partial R_k^n / \partial p_m^n$ and A_m^n are given by

$$\frac{\partial R_k^n}{\partial p_m^n} = -\frac{b\left(\gamma_k^n\right)^2}{1+\gamma_k^n} \frac{\left\| \mathbf{H}_{m,k}^n \mathbf{w}_m^n \right\|^2}{\sum_{m' \in \mathcal{M}} \beta_{m',k}^n \left\| \mathbf{H}_{m',k}^n \mathbf{w}_{m'}^n \right\|^2 p_{m'}^n}, \forall k \neq k^* \tag{4-26}$$

and

$$A_m^n = \sum_{m' \in \mathcal{M} \setminus \{m\}} \left\| \mathbf{H}_{m',k^*}^n \mathbf{w}_{m'}^n \right\|^2 p_{m'}^2 + \sigma^2 \tag{4-27}$$

respectively.

The obtained $\hat{p}_m^n$ above may not in the range of [0, S]. Therefore, the real power allocation need to be adjusted following the rule given by

$$p_m^{n(t)}=\max\left\{\min\left\{\hat{p}_m^n, S\right\}, 0\right\} \tag{4-28}$$

where t indicates the times of iteration. Lagrangian multipliers can be updated in each iteration by sub-gradient method [79] as follows,

$$\begin{aligned}\lambda_m^{(t+1)} &= \max\left\{\lambda_m^{(t)} - \upsilon_\lambda^{(t)}\left(\sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K}\beta_{m,k}^m R_k^n - C_m\right), 0\right\} \\ \mu_{n,k}^{(t+1)} &= \max\left\{\mu_{n,k}^{(t)} - \upsilon_\mu^{(t)}\left(R_{\text{thres}} - R_k^n\right), 0\right\}\end{aligned} \tag{4-29}$$

where $\upsilon_\lambda^{(t)}$ and $\upsilon_\mu^{(t)}$ and the step sizes used in the current iteration for updating λ_m and $\mu_{n,k}$.

4.5 Decentralized solution

The centralized strategy of resource allocation proposed above is processed on the CU with the global CSIs at the beginning of each TTI. It is possible that the time delay caused by processing is too long to guarantee the effectiveness of a large-scale system involving numerous TPs and UEs. An alternative method for shortening time delay is a decentralized strategy that distributes calculations to each TP instead of the CU. Under a decentralized strategy, global CSIs are shared among TPs at the first place. Then resource allocation is processed at each TP independently and simultaneously, according to the known CSIs and a given strategy. In this way, the time delay of processing can be significantly decreased even in a large-scale network. However, due to the lass of knowledge about scheduling results of other TPs, the accuracy of the decentralized is unavoidably worse than the centralized.

In this section, we propose a decentralized strategy with the similar CE-based scheduling and KKT-based power allocation to the centralized proposed above. Simulation results presented in Section 4.6 will prove that the decline of system performance under the proposed decentralized strategy is acceptable.

4.5.1 Decentralized RB Scheduling Algorithm

A decentralized RB scheduling algorithm based on CE method is proposed in this

subsection. The same to the centralized proposed in subsection 4.4, the decentralized scheduling algorithm initializes the probability distribution $\mathbf{q}^m$ at the first place. Then the iteration procedure is processed to obtain RB scheduling results. It should be noticed that the decentralized strategy cannot accurately estimate data rates of ongoing transmissions due to the lass of information about scheduling results of other TPs. Therefore, the decentralized scheduling algorithm deletes the procedure of removing unqualified transmissions (stages three of the centralized). In the simulation, we consider unqualified transmissions as failures, which waste energy and contribute nothing on data rates of the system. The overall algorithm is illustrated in Fig. 4-5. **Algorithm 4-3** summarizes the decentralized RB scheduling based on CE method.

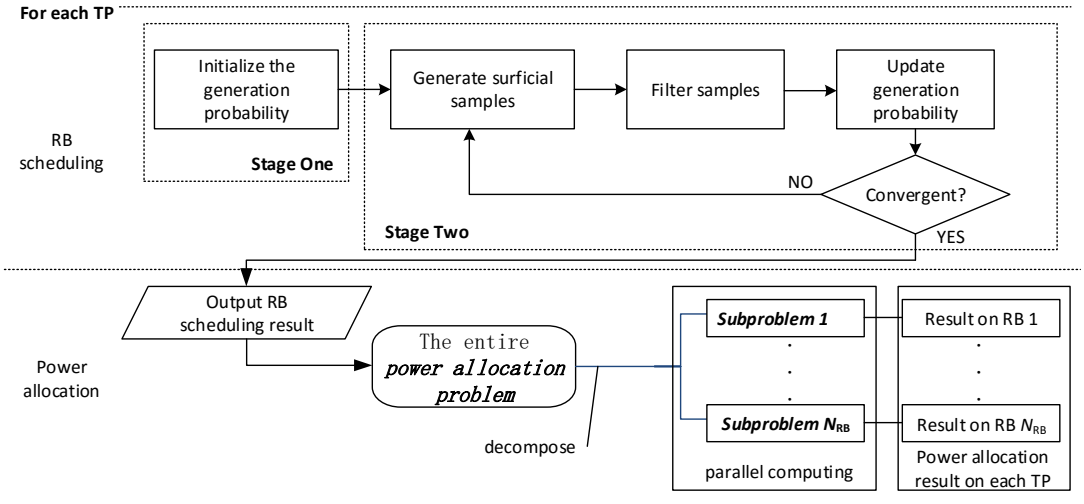


Fig.4-5 The flowchart of decentralized algorithm

Algorithm 4-3 Decentralized RB scheduling algorithm

1: Calculate R_k^n according to Eq. (4-11).

2: Initialized probability distribution $\mathbf{q}^m$ ($\forall m$) according to Eq. (4-18).

3: **for** t = 1:$t_{\max}$ **do**

4: Process **Algorithm 1**, and output $\mathbf{X}_{\text{out}}^m(\forall m)$.

5: **if** (all the elements in $\mathbf{q}^m$=[$q^{m,n}$, ($\forall n$)] converge) **then**

6: Output $\mathbf{X}_{\text{out}}$

7: **BREAK**

8: **end if**

9: **end for**

10:Map obtained $\mathbf{X}_{\text{out}}^m$ into $\left\{\beta_{m,k}^n\right\}$.

4.5.2 Power Allocation Algorithm

In this subsection, we modify the power allocation algorithm proposed in subsection 4.4.2 to be decentralized, so that it can be processed at each TP independently and simultaneously. The individual power allocation problem of each TP is given by

$$\begin{aligned} \max_{p_m^n,\forall n} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N_{\text{RB}}}\beta_{m,k}^n\frac{R_k^n}{\overline{R}_k}\Big/\sum_{k=1}^{K}\sum_{n=1}^{N_{\text{RB}}}\beta_{m,k}^n p_m^n \\ \text{s.t.} \quad & \text{C1} \quad 0\le p_m^n\le S,\forall n \\ & \text{C3} \quad \sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K}\beta_{m,k}^n R_k^n\le C_m \\ & \text{C4} \quad R_k^n\ge R_{\text{thres}},\forall n,k\in\mathcal{U}_m \end{aligned} \tag{4-30}$$

As in subsection 4.4, we constitute and solve the dual function of Eq. (4-30), instead of solving it directly. The dual function in terms of constraints C3 and C4 is given as

$$\begin{aligned} \max_{p_m^n} \quad & \sum_{n=1}^{N_{\text{RB}}}\left(\sum_{k=1}^{K}\beta_{m,k}^n\frac{R_k^n}{\overline{R}_k}\Big/\sum_{k=1}^{K}\beta_{m,k}^n p_m^n\right)-\lambda_m\left(\sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K}\beta_{m,k}^n R-C_m\right) \\ & +\sum_{n=1}^{N_{\text{RB}}}\sum_{k=1}^{K}\mu_{n,k}\beta_{m,k}^n\left(R_k^n-R_{\text{thres}}\right) \\ \text{s.t.} \quad & 0\le p_m^n\le S,\forall n \end{aligned} \tag{4-31}$$

Substituting Eq. (4-23) into Eq. (4-31), we can decompose the power allocation problem into N_{RB} independent subproblems. Let $f_{m,n}$ denote the objective of the subproblem on RB n of TP m, which is given as

$$f_{m,n}=\sum_{k=1}^{K}\beta_{m,k}^n\frac{R_k^n}{\overline{R}_k}\Big/\sum_{k=1}^{K}\beta_{m,k}^n p_m^n-\sum_{k=1}^{K}\lambda_m\beta_{m,k}^n\left(R_k^n-\frac{C_m}{\beta_m}\right)+\sum_{k=1}^{K}\mu_{n,k}\beta_{m,k}^n\left(R_k^n-R_{\text{thres}}\right) \tag{4-32}$$

$\sum_{k=1}^{K}\beta_{m,k}^n=0$ means that TP m does not schedule any transmissions on RB n. In this case, power allocation is not required, i.e., p_m^n=0. In the case where $\sum_{k=1}^{K}\beta_{m,k}^n=1$, transmit power p_m^n can be obtained by an iterative method proposed in the rest of this subsection.

Let k^* denote the UE of TP m scheduled on RB n (i.e., $\sum_{k=1}^{K}\beta_{m,k}^n=\beta_{m,k^*}^n=1$). Take the first order derivative of $f_{m,n}$ with respect to p_m^n, we obtain

$$\frac{\partial f_{m,n}}{\partial p_m^n}=\left(\frac{1}{\overline{R}_{k^*}p_m^n}-\lambda_m+\mu_{n,k^*}\right)\frac{\partial R_{k^*}^n}{\partial p_m^n}-\frac{R_{k^*}^n}{\overline{R}_{k^*}\left(p_m^n\right)^2} \tag{4-33}$$

where

$$\frac{\partial R_{k^*}^n}{\partial p_m^n}=\frac{b\sum_{m\in\mathcal{M}}\beta_{m,k^*}^n\left\|\mathbf{H}_{m,k^*}^n\mathbf{w}_m^n\right\|^2}{\left\|\mathbf{H}_{m,k^*}^n\mathbf{w}_m^n\right\|^2p_m^n+\sum_{m'\in\mathcal{M}\setminus\{m\}}\left\|\mathbf{H}_{m',k^*}^n\mathbf{w}_{m'}^n\right\|^2p_{m'}^n+\sigma^2} \tag{4-34}$$

Let $\partial f_n/\partial p_m^n$ =0. An expression of p_m^n in the tth iteration can be obtained by solving the equation, which is given as

$$p_m^{n(t)}=\sqrt{\frac{R_{k^*}^n}{\overline{R}_{k^*}}}\times\sqrt{\frac{\left\|\mathbf{H}_{m,k^*}^n\mathbf{w}_m^n\right\|^2p_m^{n(t)}+\sum_{\substack{m'\in\mathcal{M}\\m'\neq m}}\left\|\mathbf{H}_{m',k^*}^n\mathbf{w}_{m'}^n\right\|^2p_{m'}^{n(t)}+\sigma^2}{b\left\|\mathbf{H}_{m,k^*}^n\mathbf{w}_m^n\right\|^2\left(\frac{1}{\overline{R}_{k^*}p_m^{n(t)}}-\left(\lambda_m-\mu_{n,k^*}\right)\right)}} \tag{4-35}$$

where multipliers λ_m and μ_{n,k^*} should be updated according to Eq. (4-29). A suboptimal p_m^n can be approached after sufficient iterations.

4.6 Simulation and analysis

We consider a HetNet downlink system with 37 TPs, where only 19 TPs of them conduct actual communications to UEs and the others wrap them around to produce virtual interference. The radius of each small cell is 250m since a dense deployment is considered. The system includes 100 RBs, each of which has a bandwidth of 180 kHz. Therefore, the overall bandwidth of the system is 18MHz. In addition, 2×2 MIMO links are created using space channel model (SCM) [80]. Each simulation lasts 20 TTIs where a TTI is 1 ms. Important parameters used in the simulation are listed in Table 4-2.

Simulation is conducted to prove the proposed algorithms effective. A greedy algorithm, named as max capacity, is also simulated as a benchmark. Max capacity algorithm tends to allocate resources to UEs with good channel conditions, in order to achieve the optimal throughput of the network. In this way, UEs with worse channel conditions are possible to have no chance to communicate. Therefore, the fairness of max capacity algorithm is unfavorable. Figs. (4-6), (4-7) and (4-8) compare the performances of max capacity algorithm to that of the proposed.

Table 4-2 Parameters in the simulation

Parameter	Value
Layout of cells	37 hexagon cells; wrap-around used
Radius of cells	250m
Central frequency	2GHz
Number of RBs, N_{RB}	100
Limit of transmit power, S	20Watt
$N_T \times N_R$	2×2
Number of TTIs /duration of TTI	20 / 1ms
α	0.1
Channel model	SCM (pathloss + shadowing + MIMO fading)
Minimal distance (TP and UE)	35m
Height of transmit/receive antenna	35m / 1.5m
Penetration loss	20dB
Traffic model	Full buffer
Speed of UE	10m/s

Figs. 4-6 and 4-7 demonstrate average throughput per TP and fairness factor (as defined by Eq. (3-3)) of the system under different resource allocation algorithms, when transmit power of each TP is 20Watt. It is obvious that max capacity algorithm achieves an outstanding throughput, and a much worse fairness factor than the proposed. The future mobile system targets to provide not only high throughput of the network, but also quality service to every UE. Therefore, the max capacity is no longer appropriate. The proposed algorithms have much better fairness factors. More importantly, as shown in Fig. 4-8, the proposed centralized algorithm can also achieve energy efficiency as well as that of max capacity.

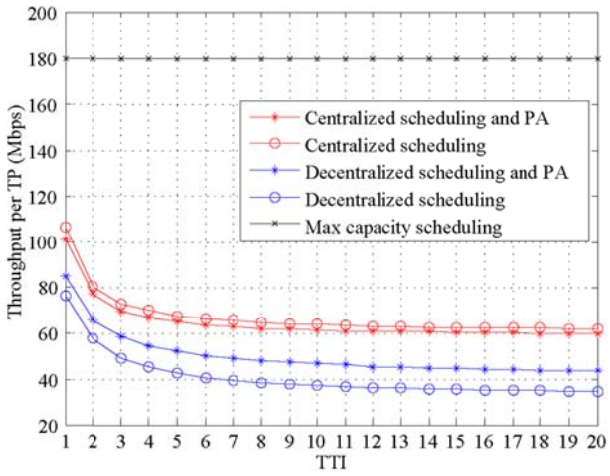


Figure 4-6 Throughput per TP under different resource allocation algorithms (20Watt, infinite C_m, R_{thres}=180kbps)

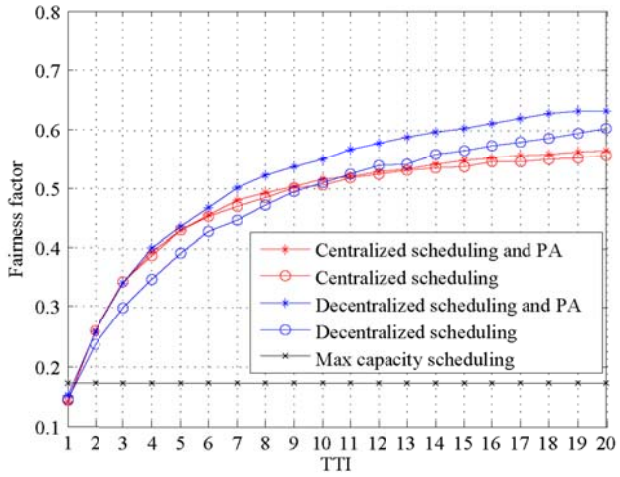


Figure 4-7 Fairness factor of UEs' data rates under different resource allocation algorithms (20Watt, infinite C_m, R_{thres}=180kbps)

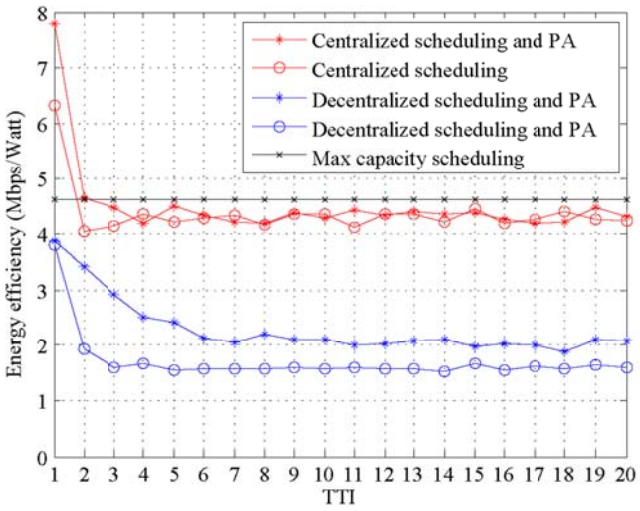
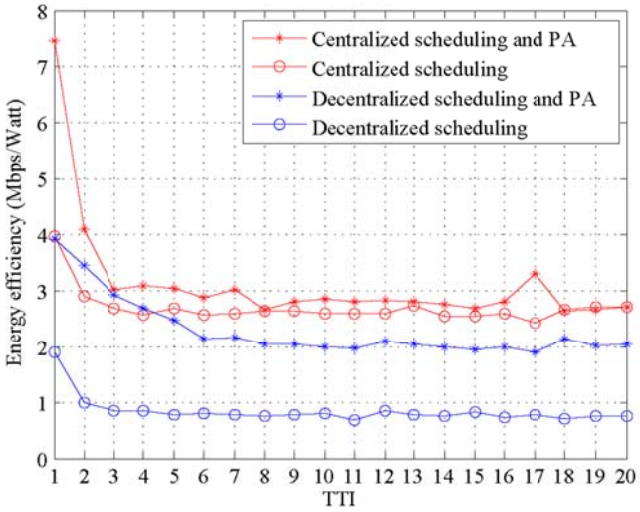


Figure 4-8 Energy Efficiencies under different resource allocation algorithms. (20Watt, infinite C_m, R_{thres}=180kbps)

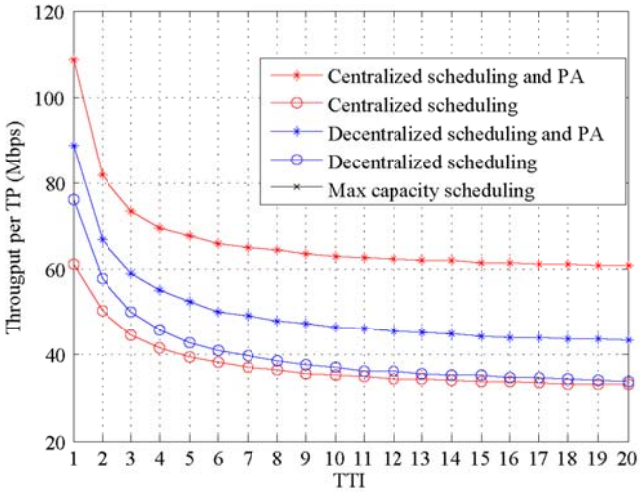
Figs. 4-6, 4-7 and 4-8 also compare the performance of the centralized algorithm proposed in Section 4.4 and that of the decentralized algorithm in Section 4.5. The results show that both energy efficiency and throughput are decreased when decentralized algorithm is used.

Fig. 4-9 shows system performances of the proposed algorithms when transmit power of each TP is 40Watt. Comparing the results to Figs. 4-6, 4-7 and 4-8, it can be seen that the energy efficiency of the centralized algorithm significantly reduces when transmit power of each TP is up to 40Watt, while the throughput does not increase. This proves that high-level transmit power is not suitable in a dense network. Additionally, the results show that the decentralized algorithm is more robust since both energy efficiency and throughput are changed a little when different transmit

power is used.



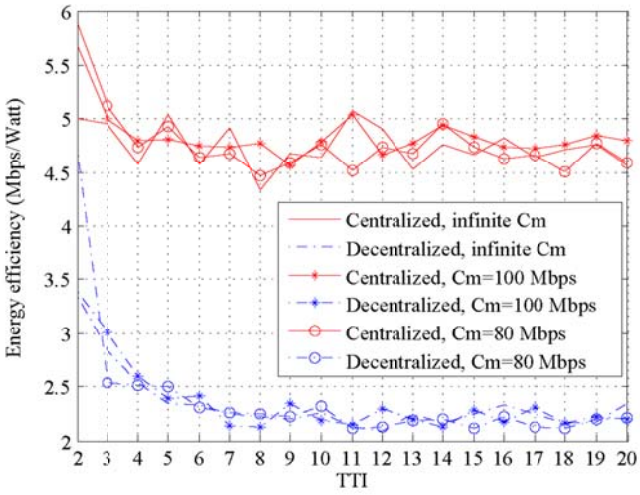
(a) Energy efficiency



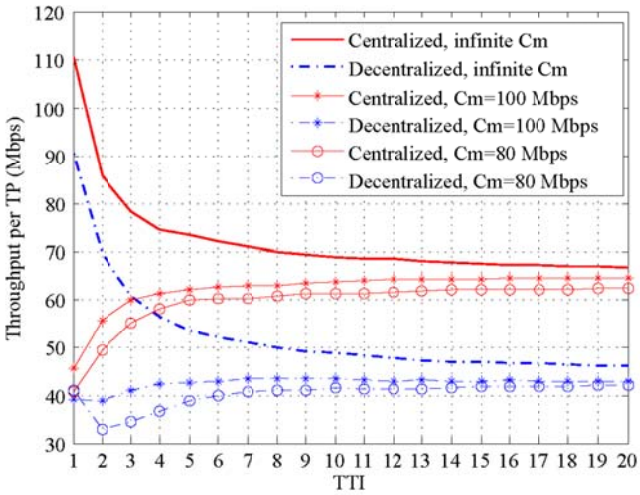
(b) Throughput per TP

Fig. 4-9 System performances under different strategies
(40Watt, infinite C_m, R_{thres}=180kbps)

Simulations are also conducted under infinite, 100 Mbps and 80 Mbps backhaul limits, respectively, with a fixed R_{thres} of 360 kbps, to clearly illustrate the effect on system performance caused by backhaul constraints. Restricted backhaul capacity leads to a low throughput per TP, as shown in Fig. 4-10(a), since fewer transmissions are scheduled in this case. However, the tendency is different in terms of energy efficiency. Fig. 4-10(b) shows that the energy efficiencies of the proposed algorithms under different backhaul limits are almost the same. This is explained by the fact that the power consumed by transmissions is also reduced when backhaul capacity is restricted.



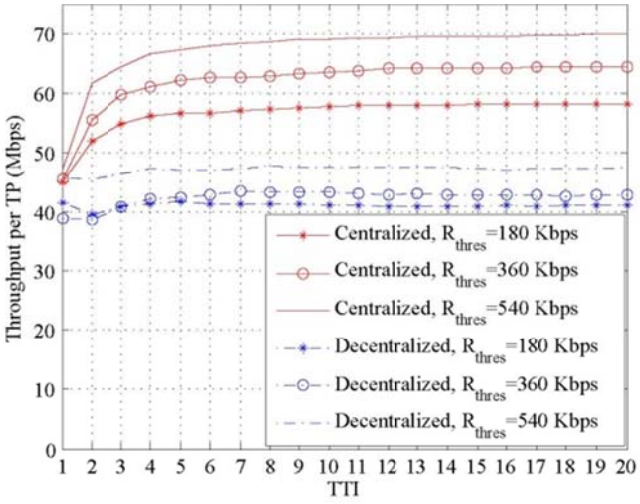
(a) Energy efficiency



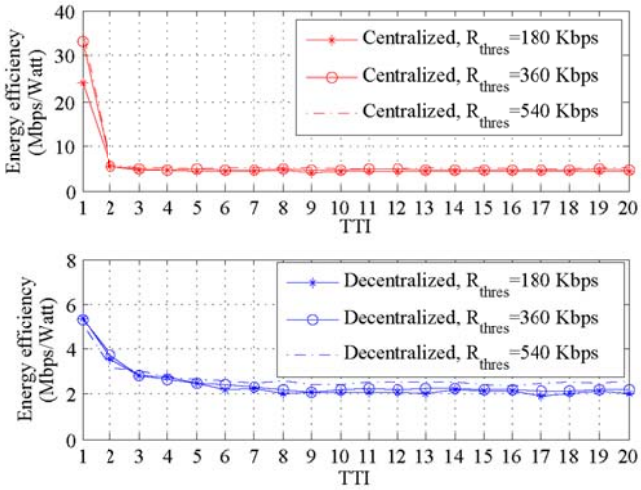
(b) Throughput per TP

Fig. 4-10 System performances under different backhaul capacity (20 Watts, R_{thres} = 360 kbps).

At last, we conduct simulations when R_{thres} is 180 kbps, 360 kbps and 540 kbps, respectively, with the fixed backhaul capacity of 100 Mbps. This shows that the throughput of our proposals grows as the increase in R_{thres}, as shown in Fig. 4-11(a). This is because more resources are assigned to the UEs with better channel conditions, which are possible to achieve for quality transmissions. Fig. 4-11(b) illustrates that the value of R_{thres} hardly affects the energy efficiency of the system. Since those transmissions estimated to be inferior to the given R_{thres} are closed, no (or little) power is wasted. Therefore, that energy efficiency of the system can be maintained at a high level.



(a) Throughput per TP



(b) Energy efficiency

Fig. 4-11 System performance under different R_{thres} (20 Watts, C_m = 100 Mbps).

4.7 Conclusion

In this chapter, we have studied a constrained RRM problem aiming at improving energy efficiency in a CoMP-based HetNet. To solve the problem, we first propose a CE-based RB scheduling algorithm under the assumption of equal power allocation. Then, a KKT-based algorithm for power allocation is presented. The proposed algorithms are considered to be used in a centralized way at the first place. Since the centralized strategy for RRM takes a long time delay in large-scale networks, we modified the proposed one in order to adapt to a decentralized system in order to shorten the time delay for processing. Simulation results compare performances of both the centralized and the decentralized and discuss the influence on system performance caused by the considered constraints.

Chapter 5 Conclusion

Spectrum and energy efficiency of the next generation wireless communication system is one of the major considerations that have obtained world-wide attention in the fields of both academic and industry. This thesis studies the RRM problems of the next generation cellular network with high technologies including CA, HetNet and CoMP, aiming at improving spectrum and energy efficiency. We first introduce the rationale behind CA, HetNet and CoMP. Then we consider a cellular network with CA and formulate the RRM involving CA allocation, RB allocation and power allocation in Chapter 3. The algorithms to solve the decomposed subproblems are also proposed. In Chapter 4, we consider a HetNet where JT CoMP is employed to reduce intercell interference. Joint spectrum and power allocation problem is studied with the objective of improving energy efficiency. In addition, in order to meet the diverse requirement in different scenarios, both centralized and decentralized algorithms for RRM are proposed. Extensive simulations are conducted to verify that the proposed algorithms are beneficial for improving resource efficiency.

Reference

[1] Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014-2019. Cisco, 2015
Available:http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking -index-vni/white_paper_c11-520862. Html

[2] 3GPP. Requirements for further advancements for Evolved Universal Terrestrial Radio Access. 3GPP, TR 36.913, 2009

[3] S. Ryoo, S. Bahk. Spatial Reuse Enhanced MAC for Wireless Dense Networks. 2009 First International Conference on Ubiquitous and Future Networks (ICUFN 2009), 225-229, 2009

[4] China Mobile. C-RAN: The Road Towards Green RAN. China Mobile, 2011
Available:http://labs.chinamobile.com/cran/wp-content/uploads/CRAN_white_paper_ v2_5_ EN(1).pdf

[5] J. G. Andrews, S. Buzzi, W. Choi, et al. What will 5G Be?. IEEE Journal on Selected Areas in Communications, 32(6), 1065-1082, 2014

[6] G. Miao, N. Himayat, A. Swami. Cross-Layer Optimization for Energy- Efficient Wireless Communications: A Survey. Wireless Communications and Mobile Computing, 9(4), 529-542, 2009

[7] Mustafa Cenk Gursoy. On the Capacity and Energy Efficiency of Training-Based Transmissions over Fading Channels. IEEE Transactions on Information Theory. 55(10), 4543-4567, 2009

[8] G. Li, H. Liu. Downlink dynamic resource allocation for multi-cell OFDMA system. 2003 IEEE 58th Vehicular Technology Conference (VTC 2003-Fall). 3, 1698-1702, 2003

[9] L. Venturino, N. Prasad, X. Wang. Coordinated Scheduling and Power Allocation in Downlink Multicell OFDMA Networks. IEEE Transactions on Vehicular Technology, 58(6), 2835-2848, 2009

[10] Y. Yu, E. Dutkiewicz, X. Huang, et al. Downlink resource allocation for next generation wireless networks with inter-cell interference. IEEE Transactions on Wireless Communications, 12(4), 783-1793, 2013

[11] Y. Wang, K. I. Pedersen, T. B. Sorensen, and P. E. Mogensen. Carrier load balancing and packet scheduling for multi-carrier systems. IEEE Transactions on Wireless

Communications, 9(5), 1780 -1789, May 2010

[12] Y. Wang, K. I. Pedersen, T. B. Sorensen, and P. E. Mogensen. Utility maximization in lte-advanced systems with carrier aggregation. 2011 IEEE 73rd Vehicular Technology Conference (VTC Spring), 1-5, May 2011

[13] F. Wu, Y. Mao, S. Leng, and X. Huang. A carrier aggregation based resource allocation scheme for pervasive wireless networks. 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing (DASC), 196 -201, Dec. 2011

[14] R. A. Abdelaal, M. H. Ismail, K. Elsayed. Resource allocation strategies based on the Signal-to-Leakage-plus-Noise Ratio in LTE-A CoMP systems. 2012 IEEE Wireless Communications and Networking Conference (WCNC), 1590-1595, 2012

[15] O. Tipmongkolsilp, S. Zaghloul, A. Jukan. The evolution of cellular backhaul technologies: current issues and future trends. IEEE Communications Surveys and Tutorials, 13(1), 97-113, 2011

[16] C. Choi, L. Scalia, T. Biermann, et al. Coordinated multipoint multiuser-MIMO transmissions over backhaul-constrained mobile access networks. 2011 IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 1336-1340, 2011

[17] A. Chowdhery, W. Yu, J. M. Cioffi. Cooperative wireless multicell OFDMA network with backhaul capacity constraints. IEEE International Conference on Communications (ICC), 1-6, 2011

[18] Q. Zhang, C. Yang, A. Molisch. Downlink Base Station Cooperative Transmission under Limited-Capacity Backhaul. IEEE Transactions on Wireless Communications, 12(8), 3746 - 3759, 2013

[19] H. Dahrouj, W. Yu. Coordinated beamforming for the multicell multi-antenna wireless system. IEEE Transactions on Wireless communications, 9(5), 1748-1759, 2010

[20] N. U. Hassan, M. Assaad. Low complexity margin adaptive resource allocation in downlink MIMO-OFDMA system. IEEE Transactions on Wireless Communications, 8(7), 3365-3371, 2009

[21] C. Hellings, M. Joham, M. Riemensberger, et al. Minimal transmit power in parallel vector broadcast channels with linear precoding. IEEE transactions on Signal Processing, 60(4), 1890-1898, 2012

[22] J. Chen, L. P. Qian, Y. Zhang. On Optimization of Joint Base Station Association and

Power Control via Benders' Decomposition. IEEE Global Telecommunications Conference, 1-6, 2009.

[23] L, P. Qian, Y. J. Zhang, Y. Wu, et al. Joint Base Station Association and Power Control via Benders' Decomposition. IEEE Transactions on Wireless Communications, 12(4), 1651-1665, 2013

[24] C. Xiong, G. Li, S. Zhang, et al. Energy-efficient resource allocation in OFDMA networks. IEEE Transactions on Communications, 60(12), 3767-3778, 2012

[25] J. B. Rao and A. O. Fapojuwo. A Survey of Energy Efficient Resource Management Techniques for Multicell Cellular Networks. IEEE Communications Surveys and Tutorials, 16(1), 154-180, 2014

[26] G. Yuan, X. Zhang, W. Wang and Y. Yang. Carrier Aggregation for LTE-Advanced Mobile Communication Systems. IEEE Communications Magazine, 48(2), 88-93, Feb. 2010

[27] 3GPP, Carrier Aggregation explained. June 2013
Available:http://www.3gpp.org/technologies/keywords-acronyms/101-carrieraggregation-explained

[28] Antitsu. Understanding LTE-Advanced Carrierd Aggregation. White Paper, Feb. 2013

[29] C. Bottai, C. Cicconetti, A. Morelli, M. Resellini, C. Vitale. Energy-efficient user association in extremely dense small cell networks. Networks and Communications (EuCNC), 1-5, 2014

[30] J. Zhang, J. Feng, C. Liu, et al. Mobility Enhancement and Performance Evaluation for 5G Ultra Dense Networks. 2015 IEEE Wireless Communications and Networking Conference (WCNC), 1793-1798, 2015

[31] W. Hashim, N.M. Anas. Power Adjustment in Dense Deployment Network an Empirical Study. TENCON 2014 - 2014 IEEE Region 10 Conference, 1-5, 2014

[32] A.G. Gotsis, S. Stefanatos, A. Alexiou. Spatial Coordination Strategies in Future Ultra Dense Wireless Networks. 2014 11th International Symposium on Wireless Communications System (ISWCS), 801-807, 2014

[33] J. Hoydis, M. Kobayashi and M. Debbah. Green Small-Cell Networks. IEEE Vehicular Technology Magazine, 6(1), 37-43, March 2011

[34] R.E. Hattachi, J. Robson. NGMN White Paper, Next Generation Mobile Networks Alliance, 2015
Available:https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0

.pdf

[35] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Lou, M. Vajapeyam, T. Yoo, Q. Song, D. Malladi. A survey on 3GPP heterogeneous networks. Wireless Communication, 18(3), 10-21, June 2011

[36] Y. Wang and K. I. Pedersen. Performance analysis of enhanced inter-cell interference coordination in LTE-Advanced heterogeneous networks. 2012 IEEE 75th Vehicular Technology Conference (VTC Spring), 1-5, 2012.

[37] K. Okino, T. Nakayama, C. Yamazaki, et al. Pico Cell Range Expansion with Interference Mitigation toward LTE-Advanced Heterogeneous Networks. 2011 IEEE International Conference on Communications Workshops (ICC), 1-5, 2011.

[38] K. Sundaresan, S. Rangarajan. Efficient Resource Management in FDMA Femto cells. In Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing (MobiHoc '09), 33-42, 2009.

[39] S. Rangan. Femto-macro cellular interference control with subband scheduling and interference cancelation. 2010 IEEE GLOBECOM Workshops (GC Wkshps), 695-700, 2010

[40] M. Peng, Y. Liu, D. Wei, W. Wang and H. H. Chen. Hierarchical cooperative relay based heterogeneous networks. Wireless Communication. 18(3), 48-56, June 2011

[41] A. Nosratinia, T.E. Hunter, A. Hedayat. Cooperative communication in wireless networks. IEEE Communications Magazine, 42(10), 74-80, 2014

[42] S. W. Halpern. Reuse partitioning in cellular systems. 33rd IEEE Conference on Vehicular Technology. 33, 322-327, 1983.

[43] 3GPP. OFDMA Downlink Inter-Cell Interference Mitigation. 3rd Generation Partnership Project (3GPP), Project Document R1-06029, 2006

[44] 3GPP. Soft Frequency Reuse Scheme for UTRAN LTE. 3rd Generation Partnership Project (3GPP), Project Document R1-050507, 2005

[45] M. Karakayal, G. Foschini and R. Valenzuela. Network coordination for spectrally efficient communications in cellular system. IEEE Wireless Communication, 13(4), 56 - 61, 2006

[46] S. Zhou, M. Zhao, X. Xu, et al. Distributed wireless communication system: a new architecture for future public wireless access. IEEE Communication Magazine, 41(3), 108 - 113, 2003

[47] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. Mayer, L.

Thiele, V. Jungnickel. Coordinated multipoint: concepts, performance, and field trial results. IEEE Communication Magazine, 2, 102-111, 2011

[48] D. Gesbert, S. Hanly, H. Huang, et al. Multi-Cell MIMO Cooperative Networks: A New Look at Interference. IEEE Journal on Selected Areas in Communications, 28(9), 1380-1408, 2010

[49] LG, Electronics. CoMP configurations and UE/eNB behaviors in LTE-Advanced. 3GPP TSG RAN 2009, R1-090213, 2009

[50] M.H.M Costa. Writing on dirty paper (Corresp.). IEEE Transactions on Information Theory, 29(3), 439-441, 1983

[51] G. Caire, S. Shamai. On achievable rates in a multi-antenna Gaussian broadcast channel. IEEE International Symposium on Information Theory Proceedings, 147, 2001

[52] Q.H. Spencer, A.L. Swindlehurst, M. Haardt Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. IEEE Transactions on Signal Processing, 52(2), 461-471, 2004

[53] P.W. Baier, M. Meurer, T. Weber, et al. Joint Transmission (JT), an alternative rationale for the downlink of time division CDMA using multi-element transmit antennas. 2000 IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications, 1-5, 2000

[54] Y. Gao, Y. Li, H. Yu, et al. Performance analysis of dynamic CoMP cell selection in LTE-advanced heterogeneous networks scenario. IEEE 2011 International Conference on Uncertainty Reasoning and Knowledge Engineering (URKE), 2, 173-176, 2011

[55] M. Feng, X. She, L. Chen, et al. Enhanced Dynamic Cell Selection with Muting Scheme for DL CoMP in LTE-A. IEEE 71st Vehicular Technology Conference, 5, 16-19, 2010

[56] S. Brueck, L. Zhao, J. Giese, et al. Centralized scheduling for joint transmission coordinated multi-point in LTE-Advanced. 2010 International ITG Workshop on Smart Antennas (WSA), 177-184, 2010

[57] T. Biermann, L. Scalia, C. Choi, et al. CoMP clustering and backhaul limitations in cooperative cellular mobile access networks. Pervasive and Mobile Computing, 8(5), 662-681, 2012

[58] A. Garcia-Armada. Snr gap approximation for m-psk-based bit loading. IEEE Transaction on Wireless Communications, 5(1), 57–60, Jan. 2006

[59] Z. Shen, J. Andrews, and B. Evans. Adaptive resource allocation in multiuser ofdm systems with proportional rate constraints. IEEE Transactions on Wireless Communications, 4(6), 2726 – 2737, Nov. 2005

[60] C. Bae and D. H. Cho. Fairness-aware adaptive resource allocation scheme in multihop ofdma systems. IEEE Communication Letters, 11(2), 134 –136, Feb. 2007.

[61] P. T. de Boer, D. Kroese, S. Mannor, and R. Rubinstein. A tutorial on the cross-entropy method. Annals of Operations Research, 134, 19-67, 2005 Available: http://dx.doi.org/10.1007/s10479-005-5724-z.

[62] J. C. Chen and C. K. Wen. A novel cognitive radio adaptation for wireless multicarrier systems. IEEE Communications Letters, 14(7), 629 –631, July 2010

[63] Y. Wang, Q. Zhang, and N. Zhang. Resource allocation based on subcarrier exchange in multiuser ofdm system. Science China Information Sciences, 56(1), 1-14, 2013

[64] W. Rhee and J. Cioffi. Increase in capacity of multiuser ofdm system using dynamic subchannel allocation. 2000 IEEE 51st Vehicular Technology Conference Proceedings (VTC 2000-Spring), 2, 1085–1089, 2000

[65] J. M. Yang, Y. P. Chen, J. T. Horng, and C. Y. Kao. Applying family competition to evolution strategies for constrained optimization. International Conference on Evolutionary Programming, 201-211, 1997

[66] M. Clerc and J. Kennedy. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. IEEE Transactions on Evolutionary Computation, 6(1), 58-73, Feb. 2002

[67] D. Gomez-Barquero, A. Bria, J. Monserrat, and N. Cardona. Minimal cost planning of dvb-h networks on existing wireless infrastructure. 2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications, 1 - 5, Sept. 2006

[68] C. Forster, I. Dickie, G. Maile, et al. Understanding the Environmental Impact of Communication Systems. Ofcom Plextek & Eftec Technical Report, 2009

[69] V. Mancuso, S. Alouf. Reducing costs and pollution in cellular networks. IEEE Communications Magazine, 49(8), 63-71, 2011

[70] Y. A. Sambo, F. Heliot, M. A. Imran. A Survey and Tutorial of Eletromagnetic Radiation and Reduction in Mobile Communication Systems. IEEE Communications Surveys and Tutorials, 17(2), 790-802, 2015

[71] O. Blume, D. Zeller, U. Barth. Approaches to Energy Efficient Wireless Access Networks. 2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP). 1-5, 2010

[72] EARTH. Driving the Energy Efficiency of Wireless Infrastructure to its Limits. EARTH, Jan. 2012

[73] Green Touch Green Meter Research Study. Reducing the Net Energy Consumption in Communications Networks by up to 90% by 2020. Green Touch White Paper, 2013

[74] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, et al. Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks. IEEE Communications Magazine, 48(11), 66-72, 2010

[75] S. Sesia, M. Baker. LTE, the UMTS Long Term Evolution - from theory to practice. WILEY, 135-137, 2011

[76] H. Maattanen, K. Hamalainen, J. Venalaiene, K. Schober, M. Enescu, M. Valkama. System-level performance of LTE-Advanced with joint transmission and dynamic point selection schemes. EURASIP Journal on Advances in Signal Processing, 2012(1), 1-18, 2012

[77] W. Yu, T. Kwon, C. Shin. Multicell coordination via joint scheduling, beamforming and power spectrum adaptation. IEEE Transaction on Wireless Communication, 12(7), 1-14, 2013

[78] R. Y. Rubinstein. Optimization of computer simulation models with rare events. European Journal of Operational Research, 99(1), 189-112, 1997

[79] D. Palomar, M. Chiang. A tutorial on decomposition methods for network utility maximization. IEEE Journal on Selected Areas in Communications, 24(8), 1439-1451, 2006

[80] J. Salo, G. D. Galdo, J. Salmi, P. Kyosti, M. Milojevic, D. Laselva, C. Schneider. MATLAB implementation of the 3GPP spatial channel model. 3GPP TR 2005, TR-25.996, 2005