

A Survey on Resource Allocation for 5G Heterogeneous Networks: Current Research, Future Trends and Challenges

Yongjun Xu ¹, Guoquan Li ¹, Miao Liu ¹, Haris Gacanin ¹, and Fumiyuki Adachi ¹

¹Affiliation not available

October 30, 2023

Abstract

In the fifth-generation (5G) wireless communications system, various service requirements under different communication environments are expected to be satisfied. As a new evolution network structure, heterogeneous networks (HetNet) have been fully studied in recent years. In contrast to conventional homogeneous networks, the key feature of HetNet is to increase the opportunity of spatial resource reuse and improve the quality of service of users by allowing small cells to cooperate in macrocell networks. However, since the mutual interference among different users and the limited resource are existing in HetNets, efficient resource allocation (RA) schemes are very important to reduce the interference and achieve spectrum sharing. In this paper, we provide a comprehensive survey on RA in HetNets for 5G communications. Specifically, we first introduce the definition and different network scenarios of HetNet. Second, RA models are discussed. Then, we present classification to analyze current RA schemes in the existing references. Finally, some challenging open issues and future research trends are addressed in this field. We also provide two effective approaches for the sixth-generation (6G) communications to solve the RA problems of future HetNets, namely, a learning-based approach and a control theory-based approach. This paper provides important information on HetNets, which can be used to guide the development of more efficient RA schemes in this area.

A Survey on Resource Allocation for 5G Heterogeneous Networks: Current Research, Future Trends and Challenges

Yongjun Xu, *Member, IEEE*, Guan Gui, *Senior Member, IEEE*, Guoquan Li, Miao Liu, *Member, IEEE*, Haris Gacanin, *Senior Member, IEEE*, and Fumiuki Adachi, *Life Fellow, IEEE*

Abstract—In the fifth-generation (5G) wireless communications system, various service requirements under different communication environments are expected to be satisfied. As a new evolution network structure, heterogeneous networks (HetNet) have been fully studied in recent years. In contrast to conventional homogeneous networks, the key feature of HetNet is to increase the opportunity of spatial resource reuse and improve the quality of service of users by allowing small cells to cooperate in macrocell networks. However, since the mutual interference among different users and the limited resource are existing in HetNets, efficient resource allocation (RA) schemes are very important to reduce the interference and achieve spectrum sharing. In this paper, we provide a comprehensive survey on RA in HetNets for 5G communications. Specifically, we first introduce the definition and different network scenarios of HetNet. Second, RA models are discussed. Then, we present classification to analyze current RA schemes in the existing references. Finally, some challenging open issues and future research trends are addressed in this field. We also provide two effective approaches for the sixth-generation (6G) communications to solve the RA problems of future HetNets, namely, a learning-based approach and a control theory-based approach. This paper provides important information on HetNets, which can be used to guide the development of more efficient RA schemes in this area.

Index Terms—Heterogeneous networks, resource allocation, spectrum efficiency, machine learning.

I. INTRODUCTION

With the exponential growth of mobile smart terminals, the fifth generation (5G) networks are designed to improve the capacity 1000 times compared from the fourth generation (4G) [1]–[6]. Moreover, spectrum efficiency (SE) improves 5 ~ 15 times compared to the 4G mobile networks [7]. The 5G mobile networks integrates various technologies, such as vehicular networking [8], device-to-device (D2D) communications [9], machine-to-machine (M2M) communications [10], Internet-of-Things (IoT) [11], cloud radio access networks (CRANs) [12], mobile edge computing (MEC) [13], cloud computing [14], unmanned aerial vehicles (UAV) [15], to make the

Y. Xu and G. Li are with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China. (E-mail: {xuyj, ligq}@cqupt.edu.cn)

G. Gui and M. Liu with College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications. (E-mail: {guiguan, lilmiao}@njupt.edu.cn)

H. Gacanin is with Nokia Bell Labs, 2018 Antwerp, Belgium (e-mail: harisg@ieee.org)

F. Adachi is the Research Organization of Electrical Communication, Tohoku University, Sendai 980-8577 Japan (e-mail: adachi@ecei.tohoku.ac.jp)

traditional communication network to realize the internet of everything [16].

From the aspect of network architecture, wireless networks evolved from homogeneous networks (HomNet) to heterogeneous networks (HetNet). 3GPP introduced HetNet in Release 12 [17], [18]. Specifically, HetNet allows many types of small-cells to coexist while overlapping macrocell networks at the same location or the same spectrum band, which greatly improves SE and decreases the coverage holes. Generally speaking, there are three different spectrum sharing strategies in HetNets [17], such as,

- **Overlay spectrum sharing:** the small-cell users (SUs) are allowed to utilize the frequency bands (FBs) if those FBs are not used by macrocell users (MUs).
- **Underlay spectrum sharing:** SUs and MUs are allowed to utilize the same FBs at the same time. However, it is necessary to effectively manage the interference to each MU receiver by a cross-tier interference power constraint.
- **Hybrid spectrum sharing:** the FBs are classified into two types: subchannels only used by SUs (i.e., support high data rates) and subchannels shared by MUs and SUs (i.e., support high spectrum utilization). That is to say that SUs with exclusive FBs are allowed to obtain good system performance by allocating more power due to no existence of the co-channel interference from MUs. In addition, the low data-rate SUs can share the FBs with MUs to support outdoor communication.

From the aspect of access modes, there are two types: open access and closed subscriber groups [19]:

- **Open access:** Under this access mode, users are allowed to associate with a small-cell base station (SBS) or a macro-cell base station (MBS) according to the coverage range. For instance, if a user terminal (UT) is within the coverage of a small-cell network, it is allowed to access it for communication preferentially. On the contrary, if a UT is out of the small cell, but it is within the macrocell network, it can associate with the MBS.
- **Closed subscriber groups:** Under this access mode, only subscribed UTs (i.e., SUs) can connect to the small cells, but the non-subscribed MUs can only access the macro-cell network whether or not they are within the coverage of small cells.

Finally, we highlight the unique features of HetNets in comparison with HomNets

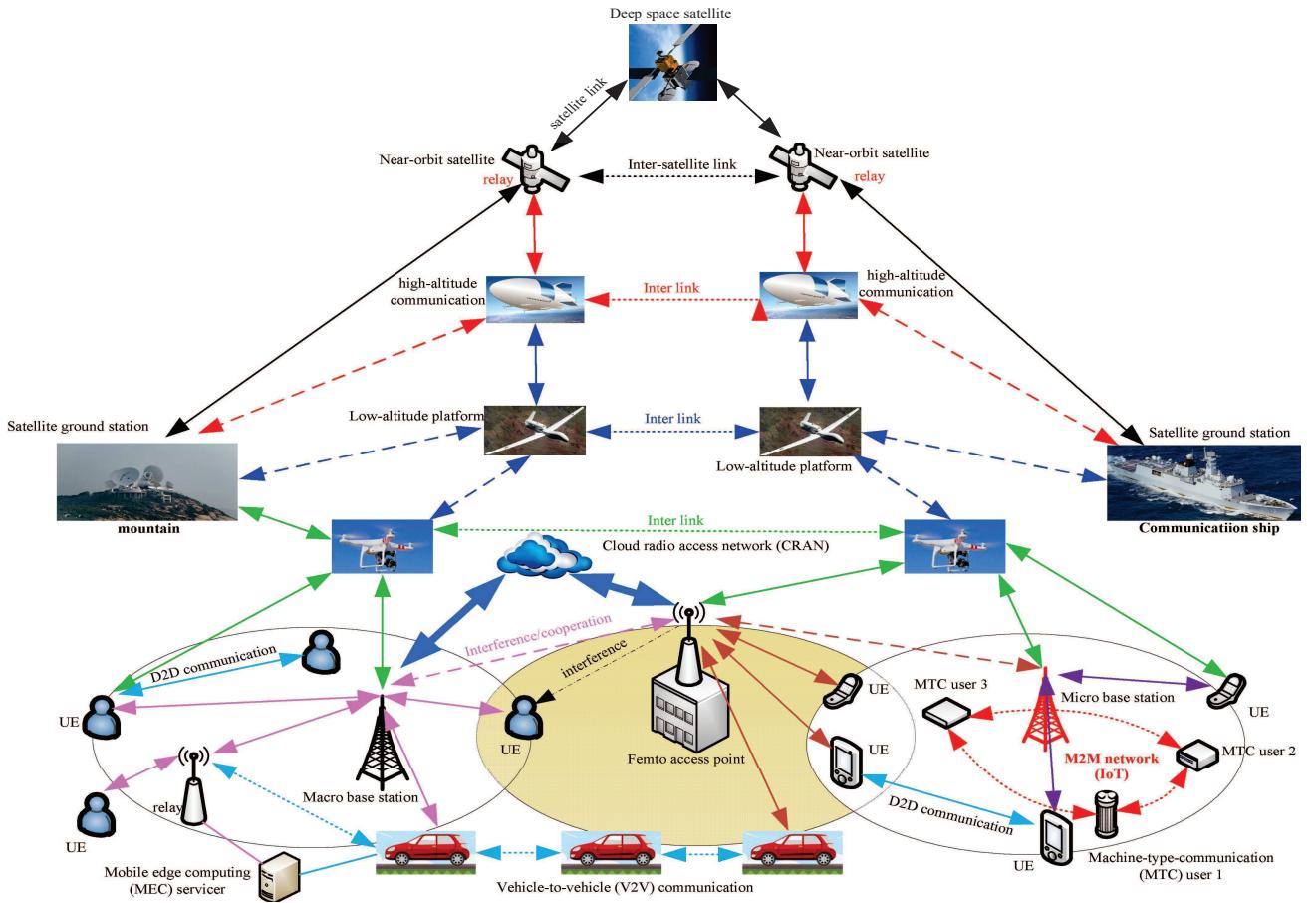


Fig. 1. System structure of future HetNets.

1) *Significant enhancement of system capacity*: more users with different access techniques are allowed to coexist in the same physical space so that the whole capacity of the communication system can be significantly improved.

2) *Ultra density*: many users with different power levels are distributed in a small range by deploying many small cells so that more users can access the network.

3) *Reduced communication blind zone*: with the distribution of different kinds of small cells (e.g., microcell, picocell, femtocell), it is possible to reduce the coverage holes and expand the communication range by placing some small access points (APs) into the poor channel environment (e.g., underground parking lots, subway).

4) *Reduced link loss and delays*: if there are no small cells in the wide-area communication scenario, the link loss or channel gains between the user equipment (UE) and the MBS is heavily degraded because of the long distance among different communication devices. If SBSs are placed between MBS and UEs, the backhaul signal from the UEs to the MBS can be achieved by a small path loss.

5) *SE improvement*: since the available SR is very limited in traditional HomNets, it is better to find an effective way to improve the SE [20]. When the transmission radius is small under the high FB, the radio frequency (RF) unit needs to be redesigned. However, HetNet significantly improves SE and provide seamless communication quality for anytime

and anywhere by coexisting different kinds of networks as indicated in Fig. 1. From the figure, diverse networks with different functions are divided into multiple tiers which cover from space to terrestrial communications.

From the perspective of interference management and resource sharing, efficient resource allocation (RA) is the key enabler of HetNet features highlighted above [21]. This paper presents a comprehensive survey on RA in HetNets. We give an overview of state-of-the-art contributions. The differences of our contributions and state of the art contributions are shown in Table I. Specifically, transmission control protocol over transmission media was summarized in [22]. A survey on secure handoff optimization schemes for multimedia services in all-IP HetNets was given in [23]. The authors in [24], [25] focused on the survey on vertical handover decision algorithms. The authors in [26] presented a survey on vehicular telematics in heterogeneous vehicular networks (HVNs), and link control, routing, congestion control, security as well as privacy were discussed in details. The authors in [27] focused on the 3GPP LTE air interface and network nodes. The authors in [28] provided the framework and challenges of HetNets from the aspect of mobile cloud computing. The authors in [29] explored a system framework of cooperative green HetNets in terms of SE and energy efficiency (EE). The authors in [30] summarized the converging solutions for heterogeneous mobile networks (HMNs). The authors in [31] focused on the

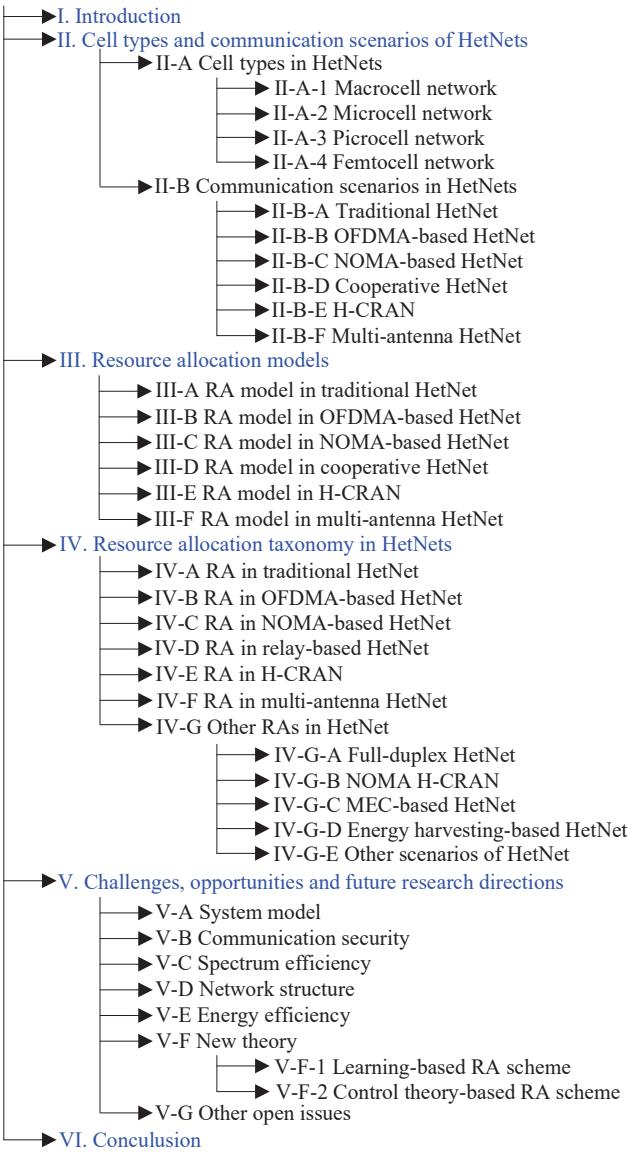


Fig. 2. The structure and contributions of this paper.

QoS and quality of experience (QoE) schemes. A survey on the HVN was presented in [32]. A brief survey and a learning approach for traffic offloading in heterogeneous cellular networks (HCNs) were introduced in [33]. The authors in [34] summarized the system architecture and key technologies of heterogeneous CRANs (H-CRANs). The authors in [35] gave the tutorials on strategies for switching off BSs. A survey about data interchange formats in the context of heterogeneous IoT (Het-IoT) was discussed in [36]. The authors investigated the survey on the Het-IoT [37].

Unlike the previously published survey papers [26]–[37], we survey both the network structures and RA models, as well as the proposed RA schemes in HetNets. This survey provides an additional value to the current works with the summary of the most recent progress on the RA problems in HetNets. In addition to the basic principle and theoretical analysis, both potential research issues and new network scenarios are presented. We discuss two comprehensive theoretical frameworks

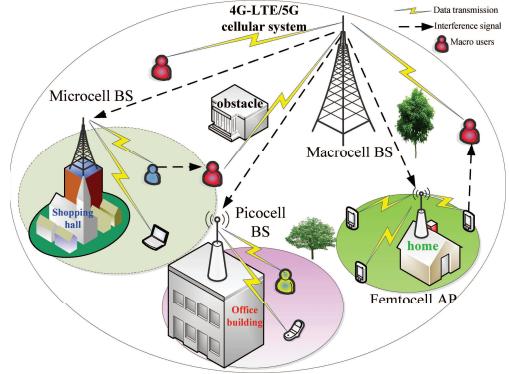


Fig. 3. System model of HetNet.

(e.g., learning-based RA and control-based RA) which are not discussed in the existing papers. For the easy understanding, Fig. 2 highlights the structure and contributions of this paper.

II. CELL TYPES AND COMMUNICATION SCENARIOS OF HETNETS

A. Cell Types of HetNets

As a potential communication way, HetNet can further improve transmission capacity and SE, as shown in Fig. 3. According to different communication ranges and application scenarios, network types can be divided into four categories as follows.

1) Macrocell networks: Macrocell network as a traditional cellular communication network can provide radio coverage by using a high-power BS, which is commonly used in current communication systems. The features of macrocell network are: (i) The MBS is always located in the high place, e.g., on the top of mountain or skyscraper, which provides a clear view over the surrounding buildings and obstacles; (ii) It has a long transmission distance and a large coverage area, where the cell radius can approach from 1 km to 25 km [38]. The distance between two neighboring MBSs is very far; (iii) Generally, the QoS of the cell-edge user is seriously affected by the shadow fading and multipath interference; (iv) There are uncovered spots or hot spots due to unevenly distributed service requests so that the QoS of the indoor users is much worse when it is served by the faraway MBS.

2) Microcell networks: Usually, microcell network is served by a low-power BS so that it is commonly used in a densely populated urban area, such as shopping hall, railway station. However, the transmission radius in the microcell network can approach only from 200 m to 1 km is smaller than that of the macrocell network. At the same time, the number of channels per network and the traffic density are dramatically increased with the decreasing frequency reuse distance of the low-power microcell BS.

3) Picocell networks: Picocell network as a small cellular network often covers a much smaller area (e.g., 100 m~200 m), such as offices, teaching buildings, subway station. Picocells are typically used to extend the coverage of indoor areas so that they can reduce the uncovered spots of indoor communication scenarios.

TABLE I
RELATED SURVEYS ON HETNETS AND COMPARISONS

Paper	Year	Network	Main addressed issues	Final target and benefits
[22]	2000	HetNet	TPC performance	Build an independent network type, to achieve TCP cope with a heterogeneous Internet via an end-to-end basis
[23]	2007	all-IP HetNet	Secure handoff optimization	Reduce the effect of handoff delay and support demanding multimedia service
[24], [25]	2009, 2010	HetNet	Vertical handover decision	Provide seamless roaming and mobility
[26]	2010	HVN	Vehicular telematics	Design the essential functional components of HVN and the protocols (radio link control, routing, congestion control, security, privacy, application development)
[27]	2011	HetNet	A comprehensive survey	Improve SE and create a network structure, overview the 3GPP LTE air interface, network nodes, cell range expansion, the enabling mechanisms in heterogeneous scenario
[28]	2013	HetNet	Mobile cloud computing	Provide vast computation resource and abundant network services
[29]	2014	HetNet	EE and SE	Cooperative HetNet to balance and optimize SE and EE.
[30]	2014	HMN	Converging solution	Improve the performance of M2M communications by using WiFi or Bluetooth
[31]	2014	HetNet	QoS and QoE mechanisms	Achieve the best possible configuration of connectivity, price and user application
[32]	2015	HVN	Architecture, challenges	Provide efficient real-time information exchange among vehicles and the wide coverage for vehicular users simultaneously
[33]	2015	HCN	EE based traffic offloading	An online reinforcement learning resolves the time-varying traffic and task offloading
[34]	2015	H-CRAN	System architecture	Fulfill the centralized cooperative process and suppress co-channel interference
[35]	2016	HetNet	BS switching	Reduce EC and meet traffic needs
[36]	2018	Het-IoT	Data exchange formats	Improve the size of transmitted messages
[37]	2018	Het-IoT	Network architecture	Achieve smart home/city, intelligent transportation, advanced manufacture, security system
This paper	2020	HetNet	RA algorithms	Achieve interference management, SR sharing, high capacity, adaptive and intelligent optimization

4) *Femtocell networks:* Femtocell (also called home BS) network has a small and low-power BS (e.g., 10 m~50 m) which is used to achieve good communication in a home place or a small business place. It is also a typical network to improve the QoS of indoor users by connecting to the home BS via wideband technology. Moreover, the installation of femtocells is much easier and more profitable than that of macrocells. In addition, femtocells can fill in the gaps of picocells and eliminate the signal loss through the buildings. The main difference between femtocells and picocells is that the number of users in the femtocells is much smaller than that of picocells, such as 4 ~ 8 users (e.g., homes) and 8 ~ 16 users per cell (e.g., small enterprise premises).

What's more, femtocells have three different access methods including open access, closed access and hybrid access [21]. The open-access mode is applied in the scenario where there is a quantity of data exchange among subnetworks, such as mall and enterprises. Under this mode, all legitimate users can access the network without additional access control. The closed access mode is mainly used in the private places, such as homes or offices, where unauthorized users cannot be allowed to access the network. Furthermore, the hybrid access mode protects the interests of legitimate users preferentially combining the above two modes. In other words, when the network is idle, users are allowed to access the network through the authentication.

Based on the above discussion, the characteristics of different networks are summarized in Table II.

B. Communication Scenarios of HetNets

Since different networks have different system models, we need to introduce the particular communication scenarios in

TABLE II
COMPARISON OF DIFFERENT NETWORK TYPES [27].

Cell types	Radius	Transmit power	Application environment
Macrocells	1km ~ 25km	20W ~ 160W	mountaintop, wide area coverage
Microcells	200m ~ 1km	2W ~ 20W	shopping hall, railway station, hot spots
Picocells	100m ~ 200m	250mW ~ 2W	office/teaching building, underground parking
Femtocells	10m ~ 50m	10mW ~ 200mW	home, small enterprise for a high data rate

HetNets before RA taxonomy. Combining the above cell types and different communication types (e.g., multi-antenna technique, cooperative communication), in this subsection, we will present several basic communication scenarios in HetNets from single channel to multiple orthogonal carriers, orthogonal multiple access (OMA) to non-orthogonal multiple access (NOMA), single-antenna system to multi-antenna system.

1) *Traditional HetNets:* The network scenario of traditional HetNet is depicted in Fig. 4. As shown in the figure, each BS and each UE has a single antenna to communicate with each other without relays. In this network structure, there are always at least two different types of cellular networks, such as macrocell and small cell (i.e., femtocell, picocell, microcell). Specifically, macrocell network as the primary network is the owner of FBs; small cell networks as the secondary network are placed in the same SR and the users control their transmit power for avoiding to significantly affect the QoS of MUs. The MBS serves multiple MUs by uplink/downlink transmission to achieve a wide coverage. However, the underlay small cells are used to achieve a higher throughput and the QoS requirements of users for certain scenarios, e.g., indoor coverage, hot spots. From the figure, there are several kinds of interference power,

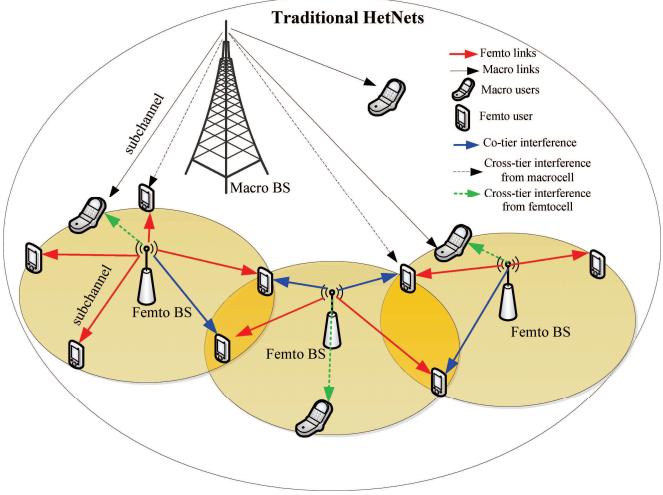


Fig. 4. Communication scenario of traditional HetNets (e.g., an example of femtocells).

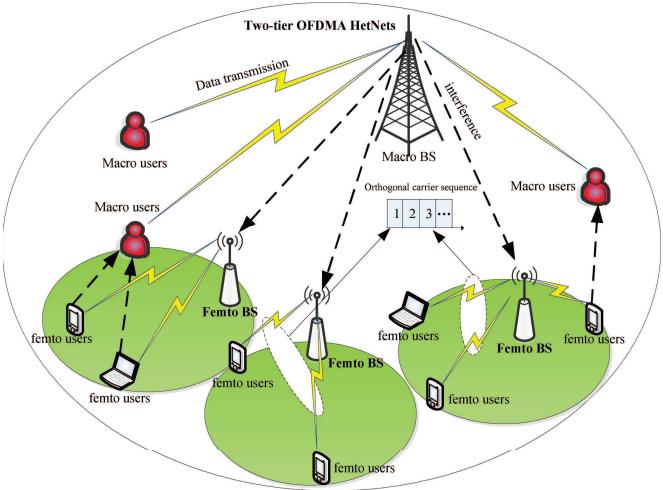


Fig. 5. Communication scenario of OFDMA-based HetNets.

such as the interference from the femto BS (FBS) to the MU receiver, the interference from the MBS to the femtocell user (FU) receiver and the interference among contiguous femtocells. Therefore, it can not only obtain a good system but also achieve the network coexistence of different cells.

2) *OFDMA based HetNets*: To reduce the mutual interference among different subchannels and improve SE, orthogonal frequency division multiplexing (OFDM) was proposed by dividing the original FB into multiple orthogonalized subcarriers in [39]. OFDM can allow flexible subcarrier allocation for realizing dynamic RA according to the demands of users. As a powerful candidate technology in HetNet, OFDMA-based HetNets have been studied by scholars. A two-tier OFDMA based HetNet is presented in Fig. 5. As shown in Fig. 5, the main difference of OFDMA based HetNets from the traditional HetNets is the use of orthogonal subcarriers (subchannels) so that there is no mutual multiple access interference (MAI) among users in the same cell.

3) *NOMA-based HetNets*: Although there is no interference among users in OMA systems (e.g., OFDMA system), the

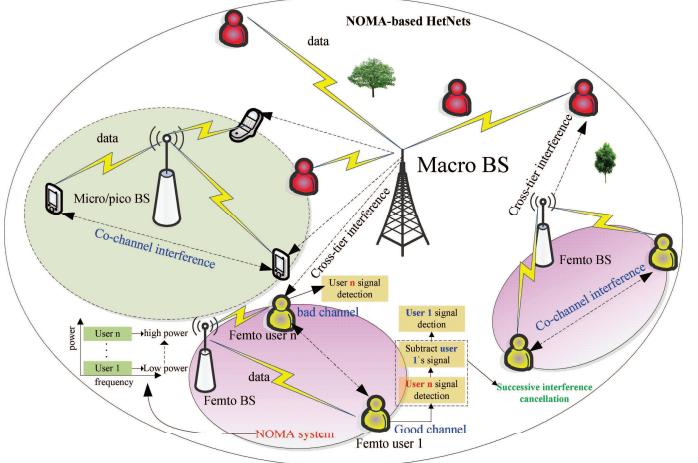


Fig. 6. Communication scenario of NOMA-based HetNets.

number of available orthogonal resources is limited. Therefore, NOMA technology was proposed to support more users for accessing networks in [43]. The basic idea of NOMA is to achieve non-orthogonal RA among different users at the cost of the increase of system increase complexity at the receiver [44]. For example, for the power-domain NOMA system, different users are allocated different power levels according to their channel quality, the same time/frequency/code resources can be shared among multiple users [45]. Meanwhile, at the receiver side, the signals of different users are decoded by using successive interference cancellation (SIC). Due to the advantage of NOMA, the NOMA technology is considered in HetNets for higher throughput and massive users access, namely, NOMA-based HetNets.

The network model of NOMA-based HetNets is given in Fig. 6. As shown in Fig. 6, there are one macrocell and multiple femtocells in the HetNet, where FUs utilize the resources by the NOMA way. The principle is: i) The FU with the poor channel condition (e.g., FU n) detects its signal by regarding the signals from other users with good channel gains as the interference power, ii) The FU with the good channel (e.g., FU 1) firstly detects the signal of the users with weak channels, and then subtracts its remodulated version from the received signal, so that the FU with the poor channel (e.g., FU n) can detect its own signal. The main difference between the NOMA system and the OMA system is that SIC is applied at the receivers.

4) *Cooperative HetNets*: Relay communication (also called cooperative communication) can effectively extend the coverage of the network. As a result, more and more researchers combine HetNets with relays to form a heterogeneous relay network for obtaining a wider coverage area and higher system throughput. A network scenario of heterogeneous relay network is presented in Fig. 7.

From this figure, different from the non-relay network, the signal transmission of this network is assisted by relays. Thus the transmission path is complex and variable, and the transmission mode is more flexible. With the introduction of relays, RA, relay selection, and forwarding methods are more

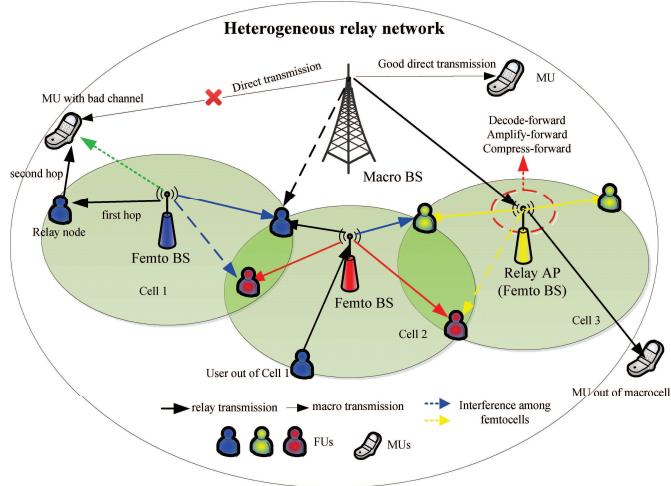


Fig. 7. Communication scenario of heterogeneous relay network.

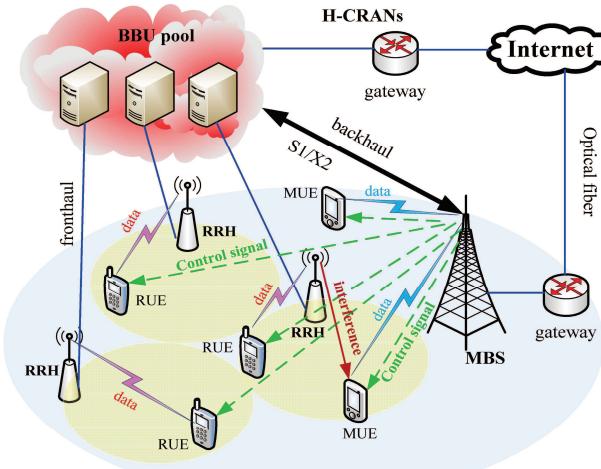


Fig. 8. Communication scenario of H-CRANs.

complex.

5) *H-CRANs*: To improve SE and reduce energy consumption (EC), H-CRAN was proposed in recent years. As a new network paradigm, H-CRAN integrating the advantages of cloud computing and HetNet can provide wider network coverage and higher throughput, and also can efficiently cope with the large-scale data processing and control. A network scenario of H-CRAN is presented in Fig. 8.

As shown in the figure, the baseband unit (BBU) pool in the cloud is used to coordinate the network resources. However, the severe inter-tier interference between the MBS and remote radio heads (RRHs) must be coordinated and eliminated for the improvement of SE and EE. The functions of different devices in H-CRANs are summarized in Table III.

6) *Multi-antenna HetNets*: Multi-antenna HetNet is another new structure of HetNets relying on spatial domain multiplexing, which can improve system capacity and SE. The MIMO channels can realize spatial division multiple access for multiple users by using transmit beamforming [40]. Different from the single-antenna channel, the direct/interfering links in multi-antenna HetNets have different characteristics. A

TABLE III
A SUMMARY OF DEVICE'S FUNCTION IN H-CRANS.

Name	Function
BBU	cloud platform, process real-time virtual computation resources manage massive signal processing and air interface protocol
RRH	wireless remote radio unit, achieve RA of users, satisfy high-speed data transmission requirements for massive data services in hotspots
MBS	achieve seamless coverage, control information transmission of the whole network, transmit control signals and system broadcast information to users, separate the function between the control plane and service plane
S1/X2	S1-communication interface between BS and evolved packet core (EPC) network, X2-interconnection interface between e-NodeBs, support the direct transmission of data and signaling

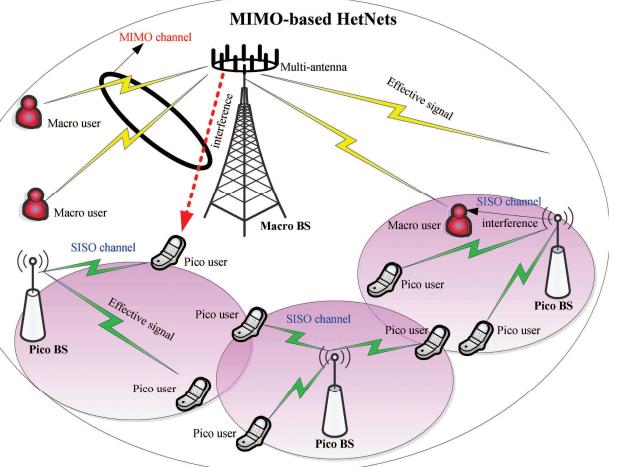


Fig. 9. Communication scenario of MIMO HetNet.

multiuser multi-antenna HetNet is presented in Fig. 9.

As shown in the figure, the multiple antennas are placed in the MBS. It can provide good communication for MUs, but causes more interference to the SUs (e.g., picocell users, PUs). Due to the heterogeneous characteristics, transmission models and interference models are different from those in the traditional MIMO-based HomNets. Moreover, there are usually three types of channels, e.g., single-input multiple-output (SIMO), multiple-input single-output (MISO) and MIMO (massive MIMO is a specifical case of the MIMO system) [42]. In order to better understand the transmission features in multi-antenna channels, the comparison of transmission signal modes is given in Fig. 10. From the figure, it is obvious that the differences among three modes are the related channel gain matrix, the dimensions of the received signal and the transmitted signals. The key issue is to design suitable beamforming vectors or precoding matrix according to channel conditions for reducing interference and improving system performance.

III. RESOURCE ALLOCATION MODELS

In order to better understand RAs, different mathematical models are introduced in this section, which is helpful to understand the differences of various RA problems, such as optimization variables (e.g., power allocation, relay selection, channel assignment).

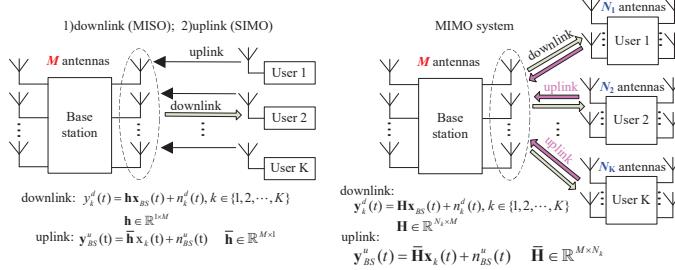


Fig. 10. Comparison of transmission signals of MIMO system.

A. RA Models in Cellular HetNets

Under this network, each user and BS have a single antenna. The BS transmits the signals to users by the FDMA way. Assume that there is one macrocell with M MUs and one small cell with N SUs under an uplink scenario, each user occupies one subchannel with bandwidth B Hz. Define the user's set as $m \in \{1, 2, \dots, M\}$ MUs and $n \in \{1, 2, \dots, N\}$ SUs. For example, the total data rate of SUs can be maximized by solving the following optimization problem

$$\begin{aligned} \max_{p_n} \sum_{n=1}^N R_n^{SU} \\ \text{s.t. } C_1 : p_n \leq P_n^{s,max} \\ C_2 : \sum_{n=1}^N p_n h_{n,m} \leq I_m^{MU}, \forall m \\ C_3 : R_n^{SU} \geq R_n^{min}, \forall n \end{aligned} \quad (1)$$

where $R_n^{SU} = B \log_2(1 + r_n)$ denotes the data rate of SU n . The SINR (r_n) is

$$r_n = \frac{p_n h_n}{\underbrace{\sum_{i=1, i \neq n}^N p_i h_{i,n}}_{\text{inter-tier interference}} + \underbrace{\sum_{m=1}^M P_m g_{m,n}}_{\text{cross-tier interference}} + \sigma_n^2} \quad (2)$$

From (2), the first item of denominator represents the interference power from other SUs' links, and the second item is the interference power from the macrocell network. h_n is the channel gain from SU n to the SBS. σ_n^2 is the background noise at SU n . P_m denotes the transmit power from MU m to the MBS. $g_{m,n}$ is the channel gain from link m to link n . $h_{i,n}$ denotes the channel gain from SU i to SU n . $P_n^{s,max}$ denotes the maximum transmit power of SU n . R_n^{min} denotes the minimum rate requirement of SU n . I_m^{MU} denotes the maximum interference power of MU m , which is used to keep the QoS requirements of MUs. It is obvious that we need to solve the optimal power p_n to maximize the sum rates of SUs.

What's more, there are other two commonly used objective functions for RA problems in such network, namely total power consumption minimization and total EE maximization. The former is always used to save energy consumption and extend network life for energy-limited networks, such as $\min \sum_{n=1}^N p_n$. The latter is used to improve unit energy utilization (i.e., bits/Joule, total rate over total energy consumption), such as $\max \frac{\sum_{n=1}^N R_n^{SU}}{\sum_{n=1}^N p_n + P_c}$, where P_c denotes the

circuit power consumption.

B. RA Models in OFDMA-based HetNets

Since the channel is divided into multiple orthogonal subcarriers (subchannels) in OFDMA-based HetNets, we firstly need to consider the subcarrier allocation problem and then determine the power allocation according to channel conditions. Assume that there are K subcarriers and $k \in \{1, 2, \dots, K\}$, we can formulate a total rate maximization problem as follows,

$$\begin{aligned} \max_{p_{n,k}, \alpha_{n,k}} \sum_{n=1}^N \sum_{k=1}^K \alpha_{n,k} R_{n,k}^{SU} \\ \text{s.t. } C_1 : \sum_{k=1}^K p_{n,k} \leq P_n^{max}, \forall n \\ C_2 : \sum_{n=1}^N \sum_{k=1}^K \alpha_{n,k} p_{n,k} h_{n,k,m} \leq I_m^{MU}, \forall m \\ C_3 : \sum_{k=1}^K \alpha_{n,k} R_{n,k}^{SU} \geq R_n^{min}, \forall n \\ C_4 : \sum_{k=1}^K \alpha_{n,k} \leq 1, \forall n \\ C_5 : \alpha_{n,k} \in \{0, 1\}, \forall n, k \end{aligned} \quad (3)$$

where $\alpha_{n,k}$ denotes the subcarrier allocation factor, which is the biggest difference with problem (1). C_4 and C_5 ensure that one subcarrier can be only used by one SU. $R_{n,k}^{SU} = \log_2(1 + r_{n,k})$ is the data rate of SU n on subcarrier k . And the SNR is

$$r_{n,k} = \frac{p_{n,k} h_{n,k}}{\sigma_{n,k}^2}, \quad (4)$$

where $p_{n,k}$, $h_{n,k}$ and $\sigma_{n,k}^2$ are the corresponding transmit power, channel gain and background noise power. Accordingly, (4) is completely different from (2). $h_{n,k,m}$ denotes the channel gain from SU n to MU m over subcarrier k .

Obviously, (3) needs to solve two variables $p_{n,k}$ and $\alpha_{n,k}$, and this problem can be extended into other scenarios with different objective functions.

C. RA Models in NOMA-based HetNets

Because multiple users can share the same subchannel under NOMA technique, the co-channel interference becomes the main challenge. For the same user scenario as Subsection A, the channel gains can be sorted as $h_1^{NOMA} \leq h_2^{NOMA} \leq \dots \leq h_n^{NOMA} \leq \dots \leq h_N^{NOMA}$. After the SIC technology, each receiver with good channel condition can perfectly decode the signals of the weakest users and then remove the inter-user interference. Thus, the transmit power satisfies $p_1^{NOMA} \geq p_2^{NOMA} \geq \dots \geq p_n^{NOMA} \geq \dots \geq p_N^{NOMA}$ [46], the SINR of SU n becomes

$$r_n^{NOMA} = \frac{p_n^{NOMA} h_n^{NOMA}}{\underbrace{\sum_{j=n+1}^N p_j^{NOMA}}_{\text{co-channel interference}} + \underbrace{\sum_{m=1}^M P_m g_{m,n}}_{\text{cross-tier interference}} + \sigma_n^2} \quad (5)$$

Other constraints are the same as problem (5).

Another important issue in RA problems of NOMA-based HetNets is the fairness among NOMA users, because the basic idea of NOMA is to improve the performance of user under bad channel condition for further capacity income [47]. As a result, the SINR of SU N with the best channel condition can be expressed as

$$r_N^{NOMA} = \frac{p_N^{NOMA} h_N^{NOMA}}{\sum_{m=1}^M P_m g_{m,N} + \sigma_N^2} \quad (6)$$

Thus the RA models of NOMA-based HetNets will become more and more complex with the increasing number of users.

D. RA Models in Cooperative HetNets

With the aid of relay nodes, RA models in cooperative HetNets are completely different from the network without relays (e.g., direct transmission). From Fig. 7, according to different relay forwarding modes, the RA models are different. Assume that there is one MBS and one FBS which serves single user, FBS acts as the relay BS due to the large-scale fading between MBS and MU. We consider an uplink transmission mode, and the MU is a source node. The destination is the MBS. Therefore, the transmission rates of source-relay (SR, $C^{S \rightarrow R}$), source-destination (SD, $C^{S \rightarrow D}$) and relay-destination (RD, $C^{R \rightarrow D}$) are

$$\begin{cases} C^{S \rightarrow R} = B \log_2 \left(1 + \frac{p^S h^{S \rightarrow R}}{\sigma^2} \right) \\ C^{S \rightarrow D} = B \log_2 \left(1 + \frac{p^S h^{S \rightarrow D}}{\sigma^2} \right) \\ C^{R \rightarrow D} = B \log_2 \left(1 + \frac{p^R h^{R \rightarrow D}}{\sigma^2} \right) \end{cases} \quad (7)$$

where p^S and p^R are the transmit power of source node and relay node. $h^{S \rightarrow D}$, $h^{S \rightarrow R}$ and $h^{R \rightarrow D}$ are the channel gains of direct transmission, SR link, and RD link. σ^2 denotes the background noise power.

And the effective data rate under the two-hop relay protocol is

$$\tilde{C}^{S \rightarrow D} = \frac{1}{2} \min \left\{ C^{S \rightarrow R}, C^{R \rightarrow D} \right\} \quad (8)$$

Therefore, there are two transmission ways. If $C^{S \rightarrow D} \geq \tilde{C}^{S \rightarrow D}$ holds, the channel gain of direct channel is better than that of relay link, so we can choose the direct transmission. Otherwise, we need to choose relay transmission. And the optimization problem can be formulated as

$$\begin{aligned} \max_{p^S, p^R} \frac{1}{2} \min & \left\{ C^{S \rightarrow R}, C^{R \rightarrow D} \right\} \\ \text{s.t. } & \hat{C}_1 : p^S \leq P_c^{\max} \\ & \hat{C}_2 : p^R \leq P_{FBS}^{\max} \end{aligned} \quad (9)$$

Obviously, (9) is completely different from RA models without relays. It also can be extended into multiple relays, cells and users under the constraints of the minimum rate requirement of each user, the allowable interference power of MUs and so on. The detailed RAAs of heterogeneous relay networks will be discussed in the following Section.

E. RA Models in H-CRANs

Comparing with traditional HetNets, the main challenge of RA models in H-CRANs is not only transmit power and

user association but also BBU offloading, resource block (RB) allocation, etc. We assume a downlink H-CRAN with one macrocell with M MUs and K RRHs with N RUEs. Each RRH is connected to the BBU pool via the wired or wireless fronthaul links. Let $\forall m \in \mathcal{R}^M = \{1, 2, \dots, M\}$, $\forall k \in \mathcal{K} = \{1, 2, \dots, K\}$ and $\forall j, n \in \mathcal{N} = \{1, 2, \dots, N\}$ denote the sets of active MUs, all RRHs and RUEs, respectively. Define the set of RRHs connected to the BBU pool via wired fronthaul links as $\mathcal{K}_1 = \{1, 2, \dots, K_1\}$, and the set of RRHs connected to the BBU pool via wireless links as $\mathcal{K}_2 = \{K_1+1, K_1+2, \dots, K\}$. If RUE n is associated in RRH k , $\alpha_{n,k} = 1$, otherwise $\alpha_{n,k} = 0$. The received SINR for RUE n accessing RRH k is given by

$$\gamma_{n,k}^R = \frac{\alpha_{n,k} p_{n,k}^R h_{n,k}^R}{\sum_k \sum_{j \neq n} \alpha_{j,k} p_{j,k}^R g_{j,n,k}^{RR} + \sum_m p_m^M g_{m,n}^{MR} + \sigma^2} \quad (10)$$

where $p_{n,k}^R$ and $h_{n,k}^R$ are the transmit power and direct channel gain of RUE n in RRH k . p_m^M denotes the allocated power from MBS to the m th MU. $g_{j,n,k}^{RR}$ is the interference channel gain of inter-tier links to RUE n . $g_{m,n}^{MR}$ is the interference channel gain of cross-tier links. σ^2 is the additive white Gaussian noise (AWGN) power.

The received SINR of MU m is given by

$$\gamma_m^M = \frac{p_m^M h_m^M}{\sum_k \sum_n \alpha_{n,k} p_{n,k}^R g_{n,k,m}^{RM} + \sigma^2} \quad (11)$$

where h_m^M is the channel gain from the MBS to MU m . $g_{n,k,m}^{RM}$ denotes the interference channel gain from link n to MU m .

Thus, based on Shannon capacity formula, the individual capacity constraint of each RRH satisfies

$$\begin{cases} \sum_n B \log(1 + \gamma_{n,k}^R) \leq R_k^1, & k \in \mathcal{K}_1 \\ \sum_n B \log(1 + \gamma_{n,k}^R) \leq R_k^2, & k \in \mathcal{K}_2 \end{cases} \quad (12)$$

where B denotes the bandwidth of each subchannel. R_k^1 and R_k^2 are the different capacity limitation of wired and wireless transmissions between the RRHs and BBU pool. Therefore, the sum capacity and total power consumption of all RRHs (C^{RRH} and P^{RRH}) are

$$C^{RRH} = \sum_k \sum_n B \log(1 + \gamma_{n,k}^R) \quad (13)$$

$$P^{RRH} = \mu \sum_k \sum_n \alpha_{n,k} p_{n,k}^R + P_c^{RRH} + P_{fh} \quad (14)$$

where μ , P_c^{RRH} and P_{fh} denote the efficiency of the power amplifier, circuit power, and power consumption of the fronthaul link, respectively. Similarly, the sum capacity and total power consumption of all MUs (C^M and P^M) are

$$C^M = \sum_m B \log(1 + \gamma_m^M) \quad (15)$$

$$P^M = \mu^M \sum_m p_m^M + P_c^M + P_{bh} \quad (16)$$

where μ^M , P_c^M and P_{bh} denote the efficiency of the power amplifier, circuit power, and power consumption of the backhaul link between the MBS and BBU pool, respectively. So, the EE maximization problem in the downlink H-CRAN can

be formulated as

$$\begin{aligned} & \max_{\{p_n^R, \alpha_{n,k}, p_m^M\}} \frac{C^{RRH} + C^M}{P^{RRH} + P^M} \\ & \text{s.t. } \tilde{C}1 : \sum_n \alpha_{n,k} = 1, \alpha_{n,k} \in \{0, 1\} \\ & \tilde{C}2 : \sum_n \alpha_{n,k} p_n^R \leq p_k^{max}, \\ & \tilde{C}3 : \gamma_{n,k}^R \geq \gamma_{n,k}^{min}, \\ & \tilde{C}4 : \gamma_m^M \geq \gamma_m^{min}, \\ & \tilde{C}5 : \sum_n B \log(1 + \gamma_{n,k}^R) \leq R_k^1, k \in \mathcal{K}_1 \\ & \tilde{C}6 : \sum_n B \log(1 + \gamma_{n,k}^R) \leq R_k^2, k \in \mathcal{K}_2 \end{aligned} \quad (17)$$

where p_k^{max} denotes the maximum transmit power of RRH k . $\gamma_{n,k}^{min}$ and γ_m^{min} present the minimum SINR requirement of each RUE and each MU, respectively. As a result, the RA in problem (17) is more complex.

F. RA Models in Multi-antenna HetNets

Comparing with the single-antenna system, multiple antennas in BSs or users make the optimal variable and channel gain of traditional RA models as multidimensional vectors. We assume a multi-antenna HetNet with one macrocell and one femtocell. The MBS with M antennas and FBS with N antennas serve multiple single-antenna users. Denote the set $\forall l \in R^M = \{1, 2, \dots, L\}$ and $\forall k, i \in R^F = \{1, 2, \dots, K\}$ as the number of MU receiver and the number of FU receiver, respectively. Hence, the transmitted signals from FBS (i.e., \mathbf{x}) can be presented as

$$\mathbf{x} = \sum_k \mathbf{w}_k s_k \quad (18)$$

where $\mathbf{w}_k \in \mathbb{R}^{N \times 1}$ and $s_k \sim CN(0, 1)$ denote the beamforming vector (i.e., precoding vector) for the k th FU and the downlink data symbol intended to the k th FU with zero mean and unit variance, respectively, and $\mathbb{E}[\|s_k\|^2] = 1$.

Because of the limitation of FBS, from (18), we have the following constraint

$$\mathbb{E}(\mathbf{x}^H \mathbf{x}) = \sum_k \|\mathbf{w}_k\|^2 \leq P_{FBS}^{max} \quad (19)$$

where $(\cdot)^H$ denotes the conjugate transpose of a matrix or vector. P_{FBS}^{max} is the maximum transmit power of FBS. Based on the same principle, we can obtain the transmitting signal and power constraint of MBS, i.e.,

$$\left\{ \begin{array}{l} \tilde{\mathbf{x}} = \sum_l \mathbf{v}_l s_l \\ E(\tilde{\mathbf{x}}^H \tilde{\mathbf{x}}) = \sum_l \|\mathbf{v}_l\|^2 \leq P_{MBS}^{max} \end{array} \right. \quad (20)$$

where $\tilde{\mathbf{x}}$ is the transmitted signal at MBS. $\mathbf{v}_l \in \mathbb{R}^{M \times 1}$ and $s_l \sim CN(0, 1)$ are the beamforming vector for the l th MU and downlink data symbol intended to the l th MU with zero mean and unit variance respectively, and $\mathbb{E}[\|s_l\|^2] = 1$. P_{MBS}^{max} is the maximum transmit power at MBS. Hence, we can easily

get the received signal at the receiver, i.e.,

$$y_k = \mathbf{h}_k^H \mathbf{w}_k s_k + \underbrace{\sum_{i \neq k} \mathbf{h}_k^H \mathbf{w}_i s_i}_{\text{inter-cell interference}} + \underbrace{\sum_l \mathbf{H}_{l,k}^H \mathbf{v}_l s_l}_{\text{MU's interference}} + n_k \quad (21)$$

$$y_l = \mathbf{H}_l^H \tilde{\mathbf{x}} + \underbrace{\sum_k \mathbf{h}_{k,l}^H \mathbf{w}_k s_k}_{\text{FU's interference}} + n_l \quad (22)$$

where y_k and y_l are the received power at FU k and MU l respectively, and $n_k \in CN(0, \sigma_k^2)$ and $n_l \in CN(0, \sigma_l^2)$ present the circularly-symmetric complex Gaussian receiver noise. $\mathbf{h}_k \in \mathbb{R}^{N \times 1}$ and $\mathbf{H}_l \in \mathbb{R}^{M \times 1}$ are the channel vectors from FBS to the k th FU and the MBS to the l th MU respectively. $\mathbf{H}_{l,k} \in \mathbb{R}^{M \times 1}$ and $\mathbf{h}_{k,l} \in \mathbb{R}^{N \times 1}$ denotes the interference channel gain vectors from macrocell network and femtocell network, respectively. Therefore, the SINR at the k th FU and interference power constraint for each MU receiver are

$$SINR_k^{\text{FU}} = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{i \neq k} |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sum_l |\mathbf{H}_{l,k}^H \mathbf{v}_l|^2 + \sigma_k^2} \quad (23)$$

$$\sum_k |\mathbf{h}_{k,l}^H \mathbf{w}_k| \leq \eta_l^{MU} \quad (24)$$

where η_l^{MU} is the permissible interference threshold for the l th MU receiver. Thus, the beamforming design with sum rate maximization of FUs can be formulated as

$$\begin{aligned} & \max_{\{\mathbf{w}_k, \mathbf{v}_l\}} \sum_k B \cdot \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{i \neq k} |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sum_l |\mathbf{H}_{l,k}^H \mathbf{v}_l|^2 + \sigma_k^2} \right) \\ & \text{s.t. } C_{1,1} : \sum_k \|\mathbf{w}_k\|^2 \leq P_{FBS}^{max} \\ & C_{2,2} : \sum_l \|\mathbf{v}_l\|^2 \leq P_{MBS}^{max} \\ & C_{3,3} : \sum_k |\mathbf{h}_{k,l}^H \mathbf{w}_k| \leq \eta_l^{MU} \end{aligned} \quad (25)$$

Obviously, it is clear that the design of beamforming vectors $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_l\}$ is crucial for the interference cancellation and optimal performance. Problem (25) can be easily extended to the other problems with different optimization objectives and MIMO-HetNets with multi-antenna BS and users. Compared with the problem (25), the difference of MIMO-HetNets is the dimension of channel gains and beamforming vectors.

IV. RA TAXONOMY IN HETNETS

In this section, we will survey RA algorithms under different network scenarios. The taxonomy for RA algorithms in HetNets is given in Fig. 11.

A. RA in Traditional HetNets

Traditional HetNets: According to Fig. 4, the key issues of RA in traditional HetNets (e.g., TDMA) are optimal power allocation (i.e., power control, RA), user association (i.e., the user uses which BS to communicate) and bandwidth allocation. The structure of this typical network is the simplest

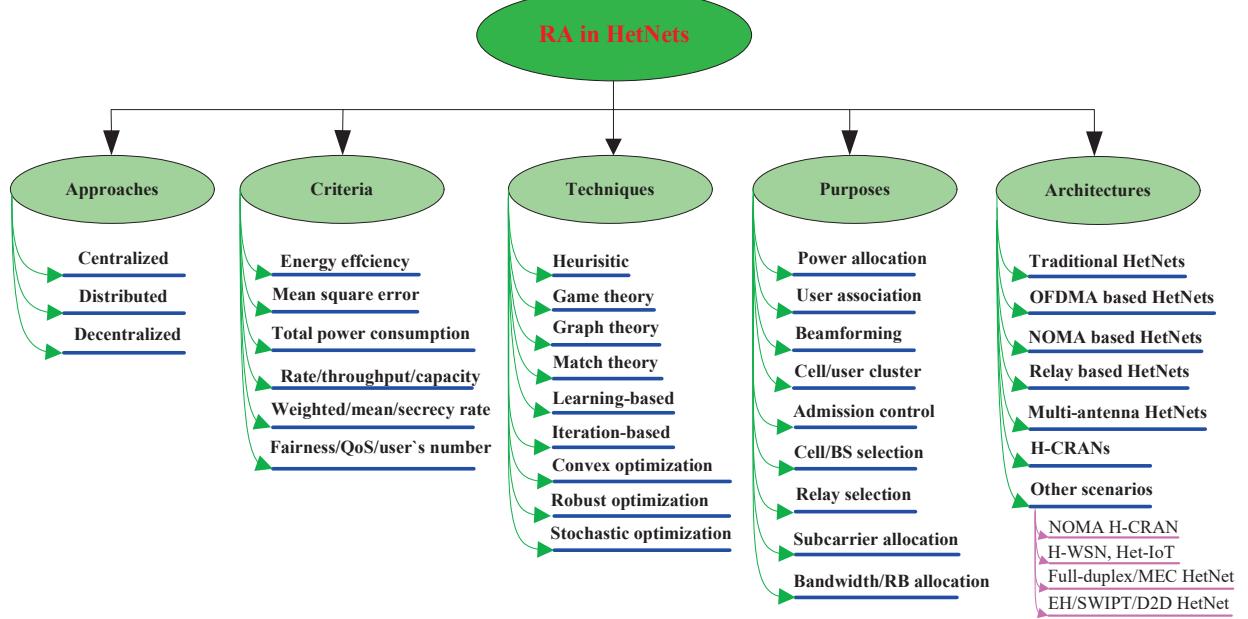


Fig. 11. Taxonomy for the RAAs in HetNets.

TABLE IV
A SUMMARY OF RA ISSUES IN TRADITIONAL HETNETS

Networks	Issues	Criteria	Description	Theory
Macro-micro [59]	Fairness [49], [63]	Mean rate [51]	Distributed algorithm [63]	Water-filling [51], Match theory [56]
Macro-small [53], [58], [63]	allocate power [48]- [63]	Secrecy rate [54]	Centralized algorithm (others)	Iteration algorithm [55], [58], [63]
Macro-femto [56], [60]-[62]	bandwidth [48], [58]	EE [52], [55], [57], [58], [60]-[62]	Imperfect CSI [59]-[61]	Convex optimization [48], [52]- [54], [59]-[62]
Macro-pic [50]-[52], [54], [55], [57]	User association [49], [51], [53], [55]	Throughput, rate [48]-[50], [56], [59], [63]	Perfect CSI (others)	Game theory [52], KKT condition [49], [54], [61], [63]

one because of no relays, antenna selection, coaching and so on. The current research works about RA problems in conventional HetNets are presented as follows.

In [48], a joint power and bandwidth allocation algorithm via convex optimization was proposed to maximize the sum throughput under the bandwidth allocation constraints, maximum transmit power constraints and the minimum data rate requirements. In [49], a fairness-driven fast RA problem for interference-free HetNets was investigated to maximize the sum of logarithms of received rates. The authors in [50] considered the optimal power allocation problem of capacity maximization by allowing each subcarrier of macrocell to be shared by users from multiple picocells. For a cellular HetNet with one macrocell and multiple picocells, the authors in [51] studied joint RA and user association problems to maximize the mean rate of the system. The authors in [52] proposed a low-complexity game-theoretic approach for EE-based RA in a two-tier HetNet. Based on fractional programming (FP), the nonconvex problem was transformed into a two-stage Stackelberg game, which is solved by using the backward induction method and the Lagrange dual decomposition (LDD) method. In [53], from the aspect of fractional frequency reuse, the authors also investigated the joint RA and user association in HetNets with one macrocell and multiple small cells. The authors in [54] studied the secrecy rate RA problem for phys-

ical layer security in HetNets with hidden eavesdropper of the macrocell. The EE-based power control and user association (i.e., BS selection, channel allocation, and model selection) problem were investigated for an uplink HetNet in [55], where the SU can associate with the BS directly or through the help of its cooperative relay. In [56], a modified many-to-one swap matching algorithm based on stable matching theory was used to solve the rate maximization RA problem. In [57], a three-stage RAA was proposed to maximize the EE of downlink transmissions in HetNets by using employ the fractional frequency reuse scheme to eliminate outages for the cell-edge users. In [58], the authors focused on EE-based maximization power allocation and wireless backhaul bandwidth allocation in downlink heterogeneous small cell networks. The authors in [59], [60] studied the robust RA problems for maximizing sum rate and EE under imperfect CSI, respectively. In [61], an uplink cross-layer RA problem under imperfect CSI was modeled as min-max fractional stochastic programming for HetNets with macrocells and femtocells, where the constraints of delay, service outage probability, system radio bandwidth, and total power consumption were considered simultaneously. In [62], a security-aware EE RA was modeled as a FP problem for HetNets where the average packet delay, the average packet dropping probability, and the total available

power consumption were considered. The authors in [63] studied the distributed RA to maximize the total throughput of the cognitive small cell networks by jointly considering interference management, fairness-based RA, average outage probability and channel reuse radius.

To give a better illustration, we give a summary of RA problems in Traditional HetNets in Table IV. *Discussion:* Based on the above-detailed discussion and Table IV, it is obvious that the user association of multicells (i.e., cell selection) is very meaningful to achieve RA problems of this network scenario. Through user association, the RA scheme can be well designed according to the user's CSI. For multiuser case, user's fairness is important for the user under the poor radio environment, however, the fairness-based RAAs are not well discussed, especially in multi-tier multicell networks with multiple users. Another question is the stability and robustness of the system. Since it is inevitable for channel perturbation, different estimation errors from channel estimation and signal's quantification and reconstruction, the trade-off between optimal performance (e.g., maximized throughput) and robustness (e.g., reduce outage event) is required to be considered ahead of time for practical communication environment. Additionally, a hardware fault tolerance based distributed RA strategy should be addressed for future research to improve data transmission efficiency (e.g., reduce the computational burden of FBS) and prolong the lifetime of femtocells.

B. RA in OFDMA-based HetNets

OFDMA-based HetNets: In OFDMA-based HetNets, the bandwidth is divided into multiple orthogonal subcarriers/subchannels. Under this orthogonal deployment, the macrocell and different small cells can be considered independently, where there is no mutual interference with each other [64]. Different from the above traditional HetNets, subchannel allocation as another key factor needs to be considered in the RA optimization problem of OFDMA-based HetNets.

In [65], the RA problem with proportional rate constraint was considered to maximize the sum rate of the system. For a two-tier downlink HetNet with multiple WLAN APs operating in OFDMA manner and one macrocell operating in TDD manner, the authors in [66] studied the EE-based RA and subcarrier allocation problem by using a double-loop iteration method. In [67], for a two-tier OFDMA heterogeneous macrocell-femtocell network, a subchannel and power allocation problem for cochannel femtocells was modeled as a mixed-integer programming problem to maximize the capacity for both delay-sensitive users and delay-tolerant users subject to the delay-sensitive users' QoS constraint and an interference constraint imposed by the macrocell. The problem was solved by using LDD method. In [68], based on Stackelberg game theory, the authors proposed an interference-aware EE RA scheme and interference pricing scheme for OFDMA-based HetNets with incomplete CSI. In [69], a joint BS selection and RA for an OFDMA HetNet with one macrocell and multiple small cells was proposed to maximize the total throughput by using game theoretic stochastic learning approaches. In [70], a joint subchannel allocation and power allocation problem was

investigated to minimize the total transmit power of BSs. The authors in [71] proposed a distributed power allocation scheme for maximizing EE in an uplink OFDMA-based HetNet where a macro-tier was augmented with small cell APs. The optimization problem was formulated as a non-cooperative game process for maximizing individual utility (i.e., its EE) of each small cell user. For a two-tier uplink HetNet with multiple small cells and one macrocell via OFDMA manner, the authors in [72] focused on the limited-feedback RAA to maximize the weighted sum rate of all users by jointly optimizing power and subcarrier allocation under the constraints of transmit power and interference power threshold. The problem was solved by using SCA and Karush-Kuhn-Tucker (KKT) conditions. For a two-tier uplink HetNet with multiple small cells and one macrocell via OFDMA manner, the authors in [73] investigated the joint power and admission control problem to maximize the SE and EE of users. With the case of local CSI overhearing and scheduling prediction, the authors in [74] studied the distributed RAA for maximizing the sum rate of FUs in an OFDMA heterogeneous macro-femto network. For a downlink OFDMA multi-cell HetNet, a joint optimization problem of user association, subchannel allocation, and power allocation was considered in [75]. The graph theory was used to solve the subproblem of user association and subchannel allocation for fixed power allocation. In addition, a KKT optimal RA solution was obtained by a low complexity algorithm based on the difference of two convex functions approximation method. In [76], the EE maximization downlink resource optimization problem was formulated as a mixed-integer nonlinear FP problem with a nonconcave nonlinear objective function and nonlinear constraints. And the problem was optimally solved by using Dinkelbach and branch-and-bound methods. In [77], for a two-tier OFDMA-based heterogeneous macro-small network with spectrum underlay and overlay access, the QoS-constrained EE optimization problem was formulated as a complex mixed-combinatorial and nonconvex optimization problem. Furthermore, with the help of appropriate decomposition, the authors proposed a dual-layer RA approach and provided a complete solution using the difference of two concave function approximations, SCA, and gradient search method.

To provide a better illustration, we give a summary of RA in OFDMA-based HetNets in Table V. *Discussion:* Based on the above introduction and Table V, it is well-known that subcarrier allocation is very important for achieving RAAs in OFDMA-based HetNets. With the introduction of the integer subcarrier assignment factor, the original RA issue becomes a mixed integer programming problem, which is converted into a continuous optimization one by aiding relaxation variables. Obviously, the problems of joint BS selection, subcarrier and robust/distributed power allocation are less addressed.

C. RA in NOMA-based HetNets

Although NOMA can bring a lot of benefits with data traffic requirements, it also brings some new challenges to the RA problems in NOMA-based HetNets due to the cross-tier interference, user's fairness, and the co-channel interference

TABLE V
A SUMMARY OF RA ISSUES IN OFDMA-BASED HETNETS

Network scenarios	one macro, multiple small cells	[64], [68], [69], [71]- [73], [76], [77]	one macro, multiple WLANs	[65], [66]
	one macro, multiple femtos	[67], [74],	one macro, multiple picos	[70]
	one macro, multiple micros	[75]		
Transmission modes	downlink	[64]- [66], [68], [70], [74]- [77]	uplink	[69], [71]- [73]
	downlink and uplink	[67]		
Purposes	power allocation	[64]- [77]	subcarrier assignment	[64], [66], [67], [70]- [72], [75]- [77]
	user association	[64], [75]	cell cluster	[64], [65]
	time fraction	[65], [66]	BS selection	[69]
Utility functions	admission control	[73]	schedule prediction	[74]
	max: weighted sum rate	[64], [72], [73], [75]	max: sum rate	[65], [67], [69], [74]
Constraints	max: EE	[66], [68], [71], [76], [77]	min: total power	[70]
	maximal transmit power	[64], [67], [72]- [77]	integer subcarrier allocation	[64], [67], [70], [72], [75], [77]
	minimum rate requirement	[65]- [67], [70], [71], [73], [76]	user association	[75]
Theory methods	MU's interference constraint	[67], [68], [72]- [74]	individual power constraint	[64]
	water-filling	[64], [74]	graph theory	[64],
	LDD	[64], [67], [72]	KKT	[65], [67], [75]
	Dinkelbach's method	[66], [76]	iteration-based method	[66]
Algorithm types	game theory	[68], [69], [71], [75]	stochastic learning	[69]
	sub-gradient method	[67], [70], [77]	SCA	[73], [77]
Algorithm types	centralized	[64]- [70], [72], [73], [75]- [77]	distributed	[71], [74]

TABLE VI
A SUMMARY OF RA ISSUES IN NOMA-BASED HETNETS

Perfect CSI	network	one macrocell, multiple picocells	[81]	one macrocell, multiple small cells	[80], [85], [86]
	transmission	one macrocell, one small cell	[83], [84]	multiple macrocells, multiple small cells	[82]
	purposes	downlink	[80]- [86]		
	utility function	power allocation	[80]- [86]	user association	[84]
		RB allocation	[80]- [86]		
		max: total EE	[82], [86]	max: fairness-based sum rate	[80], [81]
		max: sum throughput/rate	[83]- [85]		
	constraint	maximum transmit power	[80]- [83], [85], [86]	RB allocation	[80], [86]
		minimum data rate	[82], [84]- [86]	SIC constraint	[84], [85]
	theory	LDD	[86]	game theory	[80], [82], [83]
Imperfect CSI		subgradient method	[82], [86]	KKT	[80], [82], [86]
		iteration-based method	[81], [84], [85]		
	algorithm types	distributed	[83]- [85]	centralized	[80]- [82], [86]
	network	one macrocell, multiple small cells	[87], [88]	one macrocell, multiple femtocells	[89]
	transmission	downlink	[87]- [89]		
	purposes	power allocation	[87]- [89]	bandwidth allocation	[87]
	utility function	BS selection	[89]		
	constraint	max: sum rate	[87], [89]	max: EE	[88]
		maximum transmit power	[88], [89]	power allocation coefficient	[87], [88]
		minimum data rate	[89]	user association	[87], [89]
	theory	LDD	[87], [89]	SCA	[89]
	algorithm types	sub-gradient method	[87], [88]	KKT	[87]
		distributed	[87]	centralized	[88], [89]

in [78]. A survey on EE-based RA in NOMA HetNets was summarized in [79]. Based on channel conditions, RA in NOMA HetNets can be classified into two cases: perfect CSI and imperfect CSI (e.g., some errors in system parameters).

Perfect CSI: In [80], a joint spectrum allocation and power control problem was modeled as a many-to-one matching game with peer effects to achieve the sum-rate maximization and the user's proportional fairness. The sequential convex programming was used to update the power allocation. In [81], the EE and fair power allocation approach was studied for a two-tier downlink ultra-dense HetNet to improve fairness and network EE. In [82], the EE maximization power allocation approach was proposed based on a Stackelberg game. The authors in [83] studied the distributed power allocation problem for NOMA HetNets to maximize the throughput of MU and SU under the maximum transmit power constraint of each user by using a Stackelberg game. The authors in [84] studied the power allocation problem with the objective of the overall throughput maximization for a downlink NOMA HetNet where both macrocell and small cell used the NOMA approach. An iterative distributed power control algorithm was proposed to solve the RA problem. The downlink power allocation with the sum-rate maximization for a CoMP-NOMA

two-tier HetNet was considered in [85]. In [86], the problem of subchannel allocation and power allocation was formulated to maximize the overall EE of both macrocell and small cells. The convex relaxation and Lagrangian dual decomposition approaches were used to obtain a suboptimal algorithm for reducing the co-channel interference and cross-tier interference.

Imperfect CSI: The authors in [87] proposed a distributed cluster formation and power-bandwidth allocation approach for downlink HetNets with NOMA, where a non-ideal NOMA scheme was considered with power disparity and sensitivity constraints, delay tolerance, and residual interference after cancellation. The non-convex optimization problem was transformed into a convex form by using geometric programming. In [88], the authors considered the EE power allocation issue for downlink NOMA HetNets with imperfect CSI. The RA model was modeled by a probabilistic non-convex problem which was transformed into a convex problem by the sequential convex programming, and the solutions were obtained by using a bisection search algorithm. The authors in [89] proposed a downlink chance-constrained robust radio RA and BS selection algorithm for maximizing the weighted sum rate of the elastic users with channel uncertainties in a power

domain NOMA HetNet. The sample approximation scheme was used to deal with the probabilistic constraints of the user's data rate.

To provide a better illustration, we give a summary of RA in NOMA-based HetNets in Table VI. *Discussion:* According to the above discussion and analysis in Table VI, it is clear that optimal RA schemes in NOMA-based HetNets are well studied, but the distributed RA scheme with user's fairness and imperfect CSI/SIC is not considered. For the NOMA protocol, the fairness factor is a key index for good resource optimization. Additionally, SIC's residual errors and channel estimation errors are inevitable in this type of network. Low-complexity distributed RA approaches are more useful for practical system design.

D. RA in Relay-based HetNets

As we know, there are different small cells in HetNets, however, the QoS of cell-edge users and the users with the weak channels in the indoor environment may not be ensured due to the limited coverage area and weak signal strength of the BSs. As a result, the relay transmission is introduced in HetNets, also called heterogeneous relay network or heterogeneous cooperative network [90], [91]. The RA problems of relay-based HetNets are mainly concentrated in transmit power allocation and relay selection [92], since the relay selection can choose the best transmission path for good data communication. Additionally, transmit power allocation can achieve some optimization objectives, such as EE-maximization, the QoS of each user.

Single relay: The authors in [93] studied the beamforming designs of the relay user and MUE for a heterogeneous relay network, where a FU as a relay helped the uplink transmission between D2D user and the MUE. The proposed multistage maximal-ratio combining approach can make a balance between the signal of relay and MUE. In [94], the authors proposed a hybrid decode-forward compress-forward relay selection scheme and reconsidered the joint bandwidth and power allocation for a downlink heterogeneous relay network with frequency division relay channel scenario.

Multiple relays: In [95], the authors focused on the EE-based optimal relay selection and radio mode selection (i.e., multiple radio access technologies) for heterogeneous cooperative networks with the decode forward (DF) mode. In [96], a long-term proportional fair RA problem via the gradient-based method and KKT conditions was solved to maximize the sum rate of UEs for a heterogeneous relay network, where the relay nodes with in-band backhaul act as micro BSs and are able to serve UEs either independently or cooperatively with the MBS. The authors in [97] developed a hierarchical Stackelberg game to achieve mobile users' sum-rate maximization based RA for the heterogeneous relay networks. In [98], the authors studied the EE maximization RA and cell selection problems for a heterogeneous relay network with macrocells and picocells, where D2D relay nodes were used to extend the coverage of macrocells for the performance of cell-edge users. Based on the Charnes-Cooper transformation method, the original non-linear FP problem was transformed into a concave optimization problem solved by using an outer approximation

algorithm. In [99], the authors focused on the overall EC minimization of the pico-relay BSs for an overlay-based green relay assisted D2D communication scenario in HCNs. In order to avoid to predict the available green energy, the D2D users equipped with a dual battery system to harvest, store and use green energy. The overall data rate maximization based RA problem of D2D users were also studied. In [100], the authors formulated a new relay selection and power allocation problem with mixed-integer linear programming to select solar-powered relay stations and grid-powered relay stations meanwhile the optimization objective was to minimize the total grid power consumption under a DF cooperative heterogeneous cellular communication scenario. In [101], a power allocation and relay selection scheme for the underlay D2D network was designed. The idle FBS worked as a relay for the D2D transmission pair. In order to offload the distributed load from the macrocell to femtocells and reduce the resource reuse interference in HetNets, the authors in [102] investigated the relay-aided D2D based load balancing approach for multitier HetNets. Furthermore, relay selection and RA schemes were studied to minimize the potential interference and ensure the QoS of different users. In [103], the authors studied the joint RA and power control problem for a downlink cooperative D2D HetNet. The RB, power control and relay selection were considered to the total throughput maximization based RA problem which was solved by a quantum coral reefs optimization algorithm. By using a low computational complexity iterative water-filling method, the authors in [104] investigated the joint power allocation and relay selection problem for the multi-hop relay heterogeneous ultra-dense network.

To provide a better illustration, we give a summary of RA in relay-based HetNets in Table VII. *Discussion:* Based on the above discussion, relay node selection and power allocation are two important issues in relay-based HetNets. With the introduction of relays, the communication scenario becomes more complex than traditional direct transmission. The purpose of relay selection is to choose better channels for signal transmission. But relay forward ways and network structure become complex, which makes the RA problems be more difficult. Additionally, due to the successively received data with the help of relay APs, the operation status of relays is the bottleneck of relay HetNets. Security-based RA and robust RA approaches need to be given more attention in this field.

E. RA in H-CRANs

H-CRAN is a promising transmission mode for the next-generation wireless communication technique by integrating the advantages of CRNAs and HetNets. By connecting all BSs (e.g., FBS, picocell BS, PBS) of different tiers to a central processor (e.g., the cloud) through wire/wireline backhaul links, the H-CRAN can greatly provide an open, simple, controllable and flexible communication paradigm for future wireless networks [105], [106]. In the H-CRAN scenario, the high power node (always considered as MBS in H-CRAN) is used to deliver the control signals and guarantee the seamless coverage for MUs with a low data rate due to path loss. On the other side, a huge number of RRHs densely deployed in

TABLE VII
A SUMMARY OF RA ISSUES IN RELAY-BASED HETNETS

Relay scenario	Ref.	Network	Objective function	Relay strategy	Constraint	Purposes
single relay	[93]	two-hop, one user with antennas	max: SINR	AF	maximum transmit power	beamforming
	[94]	two-hop, single user	max: total rate	hybrid DF-CF		power allocation
multiple relays	[95]	two-hop, single user	max: EE	DF	end-to-end delay, maximum transmit power	time allocation, power allocation
	[96]	two-hop, multiuser	max: sum throughput		mode selection of CoMP, maximum transmit power	power allocation
	[97]		max: sum rate		RB selection, maximum transmit power	RB allocation, power allocation
	[98]		max: EE		minimum rate, maximum transmit power, BS selection	cell selection, power allocation
	[99]		min: total power		SINR constraint, maximum transmit power	power allocation
	[100]		min:total power		SINR, maximum power	relay selection, power allocation
	[101]		max: sum throughput		interference constraint, maximum transmit power	
	[102]				minimum rate, maximum transmit power	
	[103]				maximum transmit power	
	[104]	multi-hop, multiuser	max: sum rate			

the hot spots of the macrocell network, where the fronthaul links are used to connect the BBU pool and multiple RRHs. As a result, the users accessing RRH (denoted as RUEs) often have high QoS requirements with a higher priority.

The objective of RA in H-CRANs is to improve SE and EE by interference mitigation and interference suppression. To better achieve resource management, we need to find a good RAA to alleviate the burden of the BBU pool, reduce the signal overhead and severe inter-cell (or called inter-tier) interference from adjacent cells. The feature of RA in H-CRANs is necessary to consider the RB assignment and fronthaul/backhaul transmission capacity/delays.

Cellular H-CRANs: The authors in [107] studied a joint user association, power allocation, and admission control problem in a H-CRAN with the objective of overall throughput maximization. In [108], the authors focused on user association and sum-rate maximization based RA for a downlink H-CRAN with one macro RRH and several small RRHs. The problem was constrained by the minimum rate of each SU, the sum achievable rate of fronthaul links and interference protection of the MU. User association and power allocation were achieved by using the matching theory and LDD respectively. Based on the sophisticated online learning, the authors in [109] researched the EE-based maximization problem of downlink H-CRANs subject to the constraint of the number of available RBs, minimum capacity requirements of each RUE, the QoS constraint of each MU, the maximum transmit power of MBS and RRH.

OFDMA-based H-CRANs: In order to mitigate the inter-tier interference and improve EE, the authors in [110] investigated the EE optimization problem with resource assignment and power allocation for a downlink OFDMA based H-CRAN. The EE maximization problem (i.e., a kind of nonconvex FP) of overall RRHs was formulated under the constraints of the minimum data rate of each RRH, inter-tier interference from RRHs to MU's receiver, and the maximum transmit power of the RRH. The solutions of RB allocation and power allocation were achieved by using LDD methods. The authors in [111] considered the joint resource optimization and congestion control to maximize average throughput of the users serviced

by high-power nodes (HPNs) and RUEs, and balanced between throughput utility and delay performance in a downlink multiuser H-CRAN. Based on the Lyapunov optimization theory, the original stochastic optimization problem was transformed and decomposed into three convex subproblems solved by using LDD method. For an uplink OFDMA based H-CRAN, the authors in [112] investigated the EE-based RA by considering BBU offloading.

MISO H-CRANs: The authors in [113] studied the interference collaboration and beamforming design problems in a H-CRAN with one MBS and multiple RRHs to suppress the inter-tier interference. Furthermore, the expressions of overall outage probability, system sum capacity and the average bit error rate of all radio links were derived. In [114], the authors proposed a dynamic RAA for a H-CRAN with time division duplex (TDD) mode. Specifically, a clustering scheme was designed to group the RRHs into different sets; the coordinated multipoint (CoMP) communication technology was used to improve network capacity by eliminating the inter-tier interference in every set; the joint power allocation, frame structure and subcarrier selection was formulated as a mixed strategy non-cooperative game. In [115], in order to improve queue stability and achieve cooperative beamforming, the authors investigated the average weighted EE-based maximization RA optimization problem for a downlink H-CRAN with one multi-antenna MBS and multiple multi-antenna RRHs. To solve the capacity-constrained fronthaul problem, a non-convex beamformer with economical SE maximization was formulated under fronthaul capacity and transmit power constraints in [116]. Through the bisection search method, the non-convex problem was transformed into the equivalent problem solved by the weighted minimum mean square error (WMMSE) approach. By combining RRH antenna resource and BBU computation resource, the authors in [117] proposed an EE-based maximization RA scheme under the constraints on the QoS of each UE, maximum transmit power of each RRH, the fronthaul capacity and the BBU processing ability in a H-CRAN with multi-antenna RRHs. The RA problem was decomposed into a network-wide beamforming vector optimization problem and a BBU scheduling problem, which are

TABLE VIII
A SUMMARY OF RA ISSUES IN H-CRANS

Ref.	Network Types	RA Problems	Objective Function	Optimization Variables	Solutions
[107]	cellular H-CRAN (downlink, uplink)	user association, power allocation, admission control	max: throughput and the number of users	user's number, transmit power	outer approximation algorithm
[108]	cellular H-CRAN (downlink)	user association and RA	max: sum rate	RB allocation, transmit power, user association matrix	matching game
[109]	cellular H-CRAN (downlink)	green RA	max: EE of small cells	RB allocation, power allocation	sophisticated online learning
[110]	OFDMA-CRAN (downlink)	EE RA	max: EE of RRHs	subchannel allocation, power allocation	LDD, subgradient method
[111]	OFDMA H-CRAN (downlink)	EE RA, congestion control	max: EE of RRHs	traffic admission, user association, transmit power, RB allocation	Lyapunov optimization, LDD
[112]	OFDMA H-CRAN (uplink)	EE RA, BBU offloading	max: average throughput, keep network stability	AP assignment, subcarrier and power allocation, fronthaul allocation of BBUs	SCA, complementary geometric programming
[113]	MISO H-CRAN (downlink)	inter-tier interference suppression	max: throughput of RRHs	beamforming vector, outage probability analysis, power allocation	KKT conditions
[114]	MISO H-CRAN (TDD, downlink)	traffic asymmetry, inter-cell interference suppression	max: average capacity	power allocation and subcarrier selection	game theory
[115]	MISO H-CRAN (downlink)	EE RA	max: average weighted EE of RRHs	transmit power, beamforming vector	Lyapunov optimization, WMMSE approach
[116]	MISO H-CRAN (downlink)	cost-efficient RA	max: cost-coefficient EE	beamforming vector	WMMSE approach, interior point method
[117]	MISO H-CRAN (downlink)	joint BBU RRH RA	max: EE of pico RRH and macro RRH, min: the number of working BBUs	beamforming vector, user association factor, data processing rate factor	WMMSE approach, bin packing algorithm
[118]	MISO H-CRAN (downlink)	user association, power allocation	max: sun rate	user association, transmit power	variational inequality theory
[119]	MISO H-CRAN (downlink)	pilot reuse scheduling, robust beamforming	min: sum MSE, max: sum rate of RRH and MBS	beamforming vector, pilot allocation	Dsatur algorithm, convex optimization
[120]	MISO H-CRAN (downlink)	EE RA	max: EE of HPN and RRHs	beamforming vectors of RRH and HPN, power allocation	combining matrix sparseness, normalized water-filling
[121]	MIMO H-CRAN (massive, downlink)	RRH activation, robust coordinated beamforming	min: total power consumption	RRH association factor, beamforming vector	Bernstein approximation, SDP
[122]	MIMO H-CRAN (massive, downlink)	priced-based RA	max: weighted sum rate	transmit power, bandwidth allocation	convex optimization, 1-D search method

resolved by a WMMSE approach and a bin packing algorithm via the best-fit-decreasing method respectively. In [118], a centralized-distributed method via variational inequality theory was designed to achieve joint user association and power allocation in a two-tier downlink H-CRAN with one macrocell network and multiple RRHs. In order to reduce pilot consumption and the effect of incomplete CSI, a sum-rate maximization robust beamforming and pilot scheduling problem in a dense H-CRAN with one multi-antenna MBS and multiple multi-antenna RRHs was studied under maximum transmit power constraints of each RRH and the MBS in [119]. In [120], the authors addressed the EE-based RA problem by selectively cooperative transmission and power consumption model. The joint channel matrix sparseness and normalized water-filling RAAs were proposed.

Massive MIMO H-CRANs: With the consideration of imperfect CSI and power consumption of fronthaul links as well as transmit power constraint of each RRH, the authors in [121] proposed a joint RRH activation and outage constrained coordinated beamforming algorithm for MIMO H-CRANs. A conservative convex approximation was introduced by using a semidefinite program (SDP) and Bernstein-type inequality. For a mmWave massive MIMO H-CRAN, the authors in [122] investigated the bandwidth and price-based power allocation problem for maximizing the downlink weighted sum rate of the system with transmit power constraint of each RRH and fronthaul capacity constraints. The problems were solved by

the WMMSE-based iteration algorithm and the 1-D search method.

To provide a better illustration, we give a summary of RA in H-CRANs in Table VIII. *Discussion:* Currently, combining with the above discussion, RA problems in H-CRANs have been well investigated from a single-antenna network to a massive MIMO scenario. Congestion control and BBU offloading become the new challenges for resource optimization. The effect of imperfect task offloading and imperfect CSI is less considered for practical environments. Moreover, the limitation of backhaul and fronthaul ability should be constrained in RA problems.

F. RA in Multi-antenna HetNets

Different from RA problems in single-antenna HetNets (i.e., optimize transmit power p), the resource optimization problems in multi-antenna HetNets focus on the interference alignment (also called interference cancellation) and beamforming design (also called precoding method) because the transmission signal can be integrated into the beamforming vector, e.g., the following equation. Additionally, the BS with multiple antennas can automatically adjust the phase of the transmitted signal of each antenna, so that the receiver will obtain better signal gain to overcome the path loss by effectively superimposing the electromagnetic wave signal.

Traditional multi-antenna HetNets: In [123], for a MIMO HetNet with one macrocell network and one femtocell

TABLE IX
A SUMMARY OF BEAMFORMING DESIGN IN MULTI-ANTENNA HETNETS

Paper	Network	User Scenario	Performance Metrics
[123]	MIMO HetNet (one macrocell, one femtocell)	multi-antenna MBS/FBS and MUs/FUs	<i>objective:</i> throughput maximization of each FU, mean throughput analysis, precoder selection. <i>constraint:</i> MU's interference power, FBS's maximum transmit power.
[124]	MISO HetNet (multiple macrocells, multiple small cells)	multi-antenna MBSs/SBSs, single-antenna MUs/SUs	<i>objective:</i> weighted EE maximization of MUs. <i>constraint:</i> MBS's maximum power, SU's maximum interference power, each MU's instantaneous rate.
[125]	MIMO HetNet (one macrocell, multiple femtocells)	multi-antenna MBS, FBSs, users	<i>objective:</i> sum rate maximization of FUs, sum power minimization of FUs. <i>constraint:</i> FBS's maximum power, MU's interference.
[126]	MISO HetNet (one macrocell, multiple picocells)	multi-antenna MBS/PBSs, single-antenna MUs/PUs	<i>objective:</i> sum EE maximization of PUs. <i>constraint:</i> PBS's maximum transmit power, each PU's minimum rate.
[127]	MISO HetNet (one macrocell, one small cell)	multi-antenna MBS/SBS, single-antenna MU/SU	<i>objective:</i> SU's SINR maximization. <i>constraint:</i> SBS's maximum power, MU's maximum interference.
[128]	MIMO HetNet with full duplex (one macrocell, one small cell)	multi-antenna MBS/SBS, multi-antenna MUs/SU	<i>objective:</i> weighted sum rate maximization of MUs and SU's backhaul link. <i>constraint:</i> maximum transmit power of MBS and SBS.
[129]	massive MIMO HetNet (one macrocell, multiple small cells)	massive MIMO MBS with MUs, multi-antenna SBSs with SUs	<i>objective:</i> total power minimization of SUs. <i>constraint:</i> SINR outage probability of SU, maximum transmit power of each SBS.
[130]	massive MIMO HetNet (one macrocell, multiple picocells)	massive MIMO MBS with single-antenna MUs, multi-antenna PBSs with single-antenna PUs.	<i>objective:</i> weighted sum rate maximization of PUs. <i>constraint:</i> binary variable constraint of PU, maximum transmit power of PBS.
[131]	massive MIMO HetNet (one macrocell, multiple femtocells)	massive MIMO MBS with sing-antenna MUs, multi-antenna FBS with single-antenna FUs	<i>objective:</i> sum transmit power minimization of both FUs and MUs. <i>constraint:</i> SINR outage probabilities of each FU and MU.
[132]	massive MIMO HetNet (one macrocell, multiple small cells)	massive MIMO MBS with single-antenna MUs, multiple SBSs with single-antenna SUs	<i>objective:</i> sum EE maximization of MUs + SUs. <i>constraint:</i> maximum transmit power of MBS and SUs, minimum SINR requirement of each MU and SU.
[133]	massive MIMO HetNet with cluster (one macrocell, multiple small cells)	massive MIMO MBS with single-antenna MUs, multi-antenna SBSs with single-antenna SUs	<i>objective:</i> downlink sum rate maximization of SUs. <i>constraint:</i> each SBS's maximum transmit power.
[134]	massive MIMO HetNet (one macrocell, multiple small cells)	massive MIMO MBS with single-antenna MUs, single-antenna SBSs with single-antenna SUs	<i>objective:</i> weighted sum rate maximization of MUs. <i>constraint:</i> maximum transmit power of MBS.
[135]	massive MIMO HetNet with QoE aware (one macrocell, multiple small cells)	massive MIMO MBS with single-antenna MUs, multi-antenna SBSs with single-antenna SUs	<i>objective:</i> MOS maximization of SUs. <i>constraint:</i> MBS's maximum transmit power, SBS's maximum transmit power, minimum SINR and MOS constraint of each SU.
[136]	massive MIMO HetNet with cluster (one macrocell, multiple small cells)	massive MIMO MBS with single-antenna MUs, multi-antenna SBSs with single-antenna SUs	<i>objective:</i> sum rate maximization of MUs + SUs. <i>constraint:</i> minimum rate requirements of MUs and SUs, minimum backhaul rate of SBS, maximum transmit power of MBS and small cell cluster.

network, the authors addressed the downlink interference mitigation problem between FUs and MUs by maximizing the throughput of FUs under the constraints of the cross-tier interference and the required QoS. The authors in [124] investigated the EE-based coordinated beamforming design for downlink heterogeneous multicell multiuser systems under the BS with multiple antennas. The original non-convex optimization problem was converted into a polynomial form optimization problem which is solved by introducing an efficient block coordinate ascent optimization algorithm. In [125], based on an alternating optimization method, the authors tried to study the downlink beamforming design for balancing the sum-power minimization and sum-rate maximization optimization problem in a MIMO HetNet with a single macrocell and multiple femtocells. In [126], for a real-time and non-realtime heterogeneous traffic scenario, the authors proposed a two-layer EE-based RAA to jointly optimize transmit beamforming design and power allocation policies for downlink two-tier MISO HetNets comprised of a single macrocell and multiple picocells. Because the distribution function of CSI uncertainty may be difficult to obtain some times, therefore the authors in [127] studied robust beamforming design problem under the worst-case deterministic model of imperfect CSI in a two-tier MISO HetNet. In [128], to further improve capacity, the authors investigated the distributed precoding design for MIMO HetNet with full-duplex communication in wireless backhaul links. A low-complexity iterative algorithm was presented to solve the non-convex weighted sum-rate

maximization problem.

Massive MIMO HetNets: For a downlink two-tier MISO HetNet with one macrocell and multiple small cells, the authors in [129] studied the EE-based beamforming problem under macro BS with massive antennas. The beamforming optimization problem was formulated to solve the total power consumption minimization problem subject to the QoS outage probability constraint. And the problem was resolved by using Bernstein approximation and semi-definite relaxation. Considering the imperfect CSI of PBS and PU, the authors in [130] studied the power allocation and user association problem to maximize the sum rate with proportional fairness by deriving the closed-form expression of ergodic capacity under imperfect CSI in a downlink massive MIMO HetNet. The mixed-integer nonlinear programming problem with the binary variable of user association was solved by using the dual decomposition method. Most of the literature designed beamforming under the assumption of perfect CSI, but in [131], the authors studied the robust hybrid coordinated beamforming under outage probabilities of MUs and FUs for time-division duplex (TDD) massive MIMO HetNets under downlink transmission to improve the efficiency of spectrum reuse and robust against CSI uncertainty. The robust beamforming problem was formulated to minimize the total transmit power of both MUs and FUs subject to SINR outage probability constraints of each FU and MU meanwhile the problem was solved by using Bernstein-type approximation and semidefinite relaxation methods. In [132], based on the traditional zero-forcing

beamforming method, a downlink beamforming scheme for the EE-based maximization problem of both macrocell system and small cell system was proposed in a downlink MIMO HetNet scenario. For the same communication scenario as the above reference, the authors in [133] jointly considered dynamic small cell clustering and non-cooperative game-based precoding design for reducing severe interference among different small cells. The problem was formulated to maximize the sum rate of SUs under the maximum transmit power constraint. Through applying matrix stuffing and alternative direction method of multipliers, the authors in [134] investigated the fast converging robust beamforming design for the weighted sum-rate maximization problem of MUs under maximum transmit power constraint in a downlink MIMO HetNet, while imperfect CSI was considered at the transmitter. In order to enhance the user's quality of experience (QoE), the authors in [135] considered the aggregated mean opinion score (MOS) maximization-based QoE aware beamforming design problem for the downlink two-tier massive MIMO HetNet. In [136], the authors studied the small cells cluster-based RA problem in a two-tier downlink HetNet with massive MIMO, where two FBs, cellular FB, and mmWave FB were used for wireless backhaul links. The interference coordination and precoding design were formulated as a sum-rate maximization issue of MUs and SUs, which was transformed into a standard convex optimization problem solved by using an interior-point method. To provide a better illustration, we give a summary of beamforming design in multi-antenna HetNets in Table IX.

Discussion: Most of the previous works focused on downlink beamforming design in multi-antenna HetNets, however, the case of the uplink transmission scenario is less considered. The beamforming problem in the uplink mode is more challenging than that in the downlink mode and one of the reasons is that uplink is subject to the distributed power constraints. In the downlink transmission, the beamforming vectors are centralized controlled by the BS while the precoding problem in the uplink is controlled by each individual device. Moreover, the computational complexity and network expenditure are not involved in most of the works, especially in the massive MIMO scenario. For massive transmission antennas, it is better to trade-off between system cost and high performance (e.g., better diversity gain). The reason is that both EC and the hardware design complexity of RF units are the drawbacks with the increasing number of the antenna. Finally, although the case of multi-antenna BSs for beamforming design has been given more attention, the case of multi-antenna receivers needs to be more in-depth research.

G. Other RA in HetNets

Apart from the prominent RAAs in HetNets, recently a number of RA for other HetNets have also been studied due to the coming new technology, which will be discussed in this section.

1) *Full-duplex HetNet:* In [137], the authors studied the power minimization problem for a NOMA full-duplex self-hauling HetNet, and proposed an efficient iterative RA approach to avoid the backhauling bottleneck and ensure the data

rate requirement of each user. With the full-duplex technology in HCNs, the authors in [138] investigated the RAA under the gain of self-interference cancellation. Considering the same communication scenario, the authors in [139] studied the price-based power control and RA problem in full-duplex heterogeneous macrocell-femtocell networks. A triple optimization strategy was proposed for power control, subcarrier allocation, and price regulation to mitigate the cross-tier interference. Under the TDD mode, in [140], the design of the NOMA decoding order with transmit beamforming at the MBS and power allocation at SBS were jointly considered, then an iterative low-complexity RAA was developed by using the SCA and max-min method.

2) *NOMA H-CRAN:* To obtain high EE and SE as well as low-cost operation, the authors in [141] gave a short survey on the EE problem for a NOMA H-CRAN. In [142], based on the SCA and Dinkelbach method, a cross-layer EE-based RA and RRH selection algorithm for power domain NOMA H-CRANs was proposed to maximize the EE of the elastic users subject to the average delay constraint of the streaming users and the constraints, RRH selection, subcarrier, transmit power and SIC. In [143], the optimal RA scheme was studied for a cooperative NOMA HetNet.

3) *MEC-based HetNet:* In [144], to better accommodate the dramatically increasing demand for data caching and computing services, the authors studied the RA problem for information-centric virtualized HetNets with in-network caching and MEC. A distributed algorithm based on the alternating direction method of multipliers was adopted in order to solve the virtual RA problem. Thanks to the advantages of reducing task execution latency, EC of users and achieving task offloading in MEC, the authors in [145] jointly considered the radio and computational RA problem for NOMA-based MEC in HetNets to minimize the EC of all users with the task execution latency constraint and the maximum transmit power constraint of each user. The uplink power allocation problem was resolved by using sequential convex programming. To reduce the end-to-end delay of mobile service delivery and improve the user experience, in [146], the adaptive boundary algorithm of indoor small cell BSs and power optimization schemes were proposed to obtain a bigger coverage ratio of the BS for self-organizing MEC-based heterogeneous small cell networks.

4) *Energy harvesting-based HetNet:* In [147], a distributed RAA about the EE maximization was designed to achieve the optimal user association and power control for a NOMA HetNet where the BS was powered by both renewable energy harvesting (EH) and conventional grid energy. In [148], the authors studied the subchannel allocation and power control for maximizing the EE by using a low complexity subchannel matching algorithm and Lagrange dual method in a NOMA HetNet with EH. For a power-domain NOMA HetNet with EH at the UE, the authors in [149] proposed joint subcarrier and power allocation algorithms to achieve EE-based RA for each fairness method and SIC ordering. In [150], an EE-based RAA was studied for EH based D2D HetNets by using the Dinkelbach and LDD method. In [151], the authors aimed to maximize the sum-rate of the EH aided D2D links in

a two-tier HetNet by superimposing their messages on the downlink resources of mobile users, which is achieved without unduly degrading MU's throughput. In [152], for EH aided heterogeneous cognitive radio sensor networks, the RAA was proposed to achieve the sustainability of spectrum sensors and conserve the energy of data sensors. In [153], based on a non-cooperative game-theoretic approach, the authors investigated the problem of power allocation and subchannel assignment for the simultaneous wireless information and power transfer (SWIPT) enabled HetNets with the consideration of cross-tier/co-tier interference mitigation and incomplete CSI. Based on SCA methods, an iterative algorithm was used to solve the non-convex optimization problem. In order to provide a cost-effective and long-lasting power supply for energy-constrained mobile devices in HetNets, in [154], the authors investigated the EE maximization based beamforming problem for a downlink MISO HetNet with SWIPT. The problem formulation was presented to maximize the information transmission efficiency of information decoding FUs and EH efficiency of EH FUs meanwhile the beamformers were obtained by the zero forcing and mixed beamforming schemes. The authors in [155] studied the EE optimization for CoMP SWIPT HetNets meanwhile satisfying certain QoS requirements in regard to transmission rate and EH at both the macro cell and small cells.

5) Other scenarios: To overcome the limitations of limited coverage, strictly line-of-sight transmission, and mobility robustness in the visible light communication (VLC), the authors in [156] focused on the energy-aware design of network selection and RA for a HetNet combining with RF and VLC APs. In [157], a security-aware joint power and subchannel allocation problem based on the inter-network cooperation was investigated for a cognitive HetNet under imperfect spectrum sensing. The authors in [158] focused on RA for D2D communications in multi-cell multi-band HCNs. The optimization problem was formulated as the D2D communication spectrum RA among multiple microwave bands and multiple mm-wave bands in HCNs. The authors in [159] studied the robust RA problem with chance constraints to improve the throughput and reliability of NOMA HVNs where the chance constraints with channel estimation errors were converted into the deterministic ones by using the approximation of non-central Chi-square distribution. To improve SE and EE due to the large number of connectivity demands in IoT, an EE-based RA problem with imperfect SIC was studied for the NOMA heterogeneous IoT in [160]. A deep recurrent neural network based optimal RAA was proposed to reduce the computational complexity of RA. The relationships among the above works is summarized in Fig. 12.

V. CHALLENGES, OPPORTUNITIES AND FUTURE RESEARCH DIRECTIONS

According to the above content, we have discussed current research works of RA over heterogeneous multi-tier networks in detail. We identify key research challenges and directions. Our recommendations are summarized as follows.

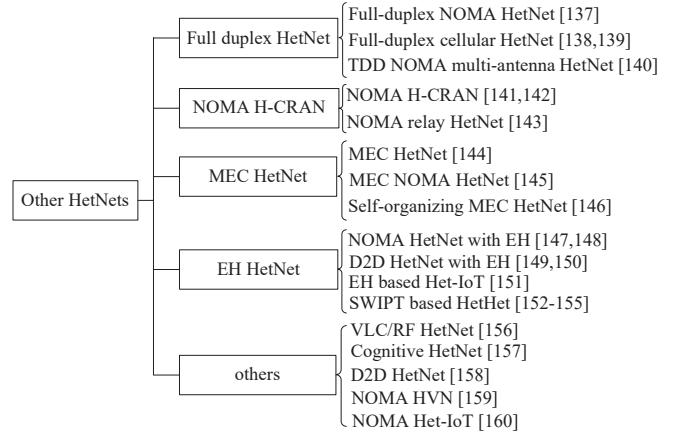


Fig. 12. The relationship on RA of other HetNets.

A. System Model

Currently, information communication technology presents a state of rapid development so that there are a lot of new technologies, such as carrier aggregation, OFDMA/NOMA, cooperation communication, massive MIMO, mmWave communication and so on. These excellent technologies would be properly integrated with the existing HetNets for better communication performance. In addition, traditional switching management methods via user, action and network status may not satisfy the QoS requirements of different types of users. In the new communication age, we not only need to consider the utilization rate of the SR but also consider the effect of channel fading and time-varying interference to the RA problems of the whole network. How to design a low complexity and high flexibility network model as well as RAAs with the consideration of user's fairness and network performance is an important research topic.

B. Communication Security

From the perspective of information transmission, information security is very important for a communication system, especially in HCNs. Although HetNets can achieve multi-network integration and satisfy different user's requirements, there may be information leakage, eavesdropping situation and other security problems. As a result, RA with the consideration of security constraints is necessary to be considered in multiuser HetNets due to complex communication scenarios, such as RA for physical-layer security in HetNets.

C. Spectrum Efficiency

As we know, the spectral resource (SR) is a precious and limited resource for wireless communication so that we need to design some methods for improving SE. In HetNets, different users in cellular networks operate in the same FBs, which cause more interference between each other. For next-generation HetNets, it moves towards the intelligent and adaptive regulation communication system. Cognitive radio technology can effectively improve the SE of the secondary market in an adaptive way, where users with cognitive function

can dynamically detect idle spectrum and use the SR by using some access methods (e.g., underlay, overlay, interweave). Therefore, the number of accessing users and throughput of the whole HetNets increase in an exponential way. Spectrum sensing-based HetNets (i.e., cognitive HetNets) is a key technique for the next generation communication system. How to tradeoff the spectrum detection capacity and dynamic RA is a very important issue in resource sharing in cognitive HetNets. The false probability may bring some new challenges for RA in different cellular networks.

D. Network Structure

From the aspect of network structure, HetNets move towards the development of better transmission efficiency, higher data rate, more powerful function and so on. As a result, how to achieve joint antenna selection and multi-user diversity optimization for multiuser MIMO HetNets is a challenging problem, since both the multi-antenna system and ultra-intensive users are a development trend. Furthermore, HetNet will move towards ground communication to space communication or underwater communication, as shown in Fig. 1. Otherwise, channel uncertainty, unreliable feedback channel (e.g., the effect of channel delay), the limited bandwidth and the scarce SR bring a lot of challenges for the design and practical application of RA in HetNets.

E. Energy Efficiency

In future communication, the main issue is to solve the operation lifetime of network and save EC for achieving green communication and reducing carbon dioxide emissions. Since there are a lot of users and small cells in multi-tier HetNets, EC is a big problem for next-generation communication. For example, RA problems for green communication-based HetNets and EH based HetNets are worth studying in our future research.

F. New Theory

1) *Learning-based RA Scheme*: With the increasing number of intelligent terminals, the BSs need to process more data than before. However, the randomness of channels and the mobility of the user as well as high-definition video services can heavily influence the performance of the communication system. As a result, we need to focus on intelligent algorithms to deal with those problems. With the development of machine learning, deep learning and artificial intelligence, many learning-based algorithms have been proposed to solve channel estimation [161]–[170], automatic modulation recognition [171]–[175], network traffic prediction [176]–[184], computing offloading [185]–[192], physical wireless techniques [193]–[203], congestion control [204]–[208], direction-of-arrival estimation [209]–[216], and so on, but there is no learning-based scheme to resolve RA problems in HetNets. Therefore, it is meaningful to design it for achieving intelligent communication.

We design a new learning-based structure to solve the RA issue in HetNets, as shown in Fig. 13. The work process is:

- (i) *CSI feedback*: Both of MU's receiver and SU's receiver feedback forward channel gains, interference power from

SUs or MUs, background noise by estimation algorithms. Also, the actual SINR and the QoS requirements of users (e.g, minimum SINR or data rate) are sent to the transmitters (MBS and SBS).

- (ii) *Optimal transmit power*: The related channel and interference information are transmitted from a local server to a cloud server by the optical fiber. The cloud server deals with massive data processing and computation. The optimal transmit power is firstly obtained by the embedded CVX optimization tool [217]. The cloud server sends the optimal value to local servers. Also, the centralized cloud server can achieve task offloading according to the requirements of MBS and SBS, which can reduce the computation burden.
- (iii) *Neural network training*: At the BS server, the CSI, interference constraint and optimization objective are considered as the inputs of neural network (e.g., a kind of learning method). Through several network training, we can obtain the actual output of the network at the output layer of the neural network. The actual deviation can be obtained by a comparator (i.e., the BS server sends another reference value). The error is feedback to the neural network for adjusting network deviation and weight updating. When the error is zero, the system is stable and the cloud server stops working. The mature networks can replace traditional algorithms (i.e., iteration-based algorithm or CVX tool) for intelligent adjustment and RA.

2) *Control theory-based RA Scheme*: From the perspective of infrastructure cost, building a large number of servers is a less realistic thing sometimes, e.g., industrial IoT, Ad-hoc network, D2D network, M2M, WSN and so on. Therefore, it is necessary to develop a new theory to effectively solve traditional resource optimization problems without increasing infrastructures. Currently, most of the RA problem is formulated as an optimization problem for achieving network requirements (e.g., throughput/EE maximization) under some constraints. There are a lot of difficulties in transformation and solution, especially in non-convex problems. Moreover, there is no general method to deal with it. As a result, we provide a new double closed loop-based RA method for better solving the resource optimization problems in HetNets, as shown in Fig. 14. The advantages of this method are summarized as follows.

- (i) *Macrocell network*: In HetNets, MUs can achieve optimization objectives by themselves without any limitations except the processing ability of MBS. As shown in Fig. 14, it can well obtain SINR-tracking performance for each user by designing a power controller via control theory, such as log-linear model [218], Fuzzy logic control [219] and robust control [220].
- (ii) *Small cell network*: Usually, small cell network is designed in the hot spots of the macrocell network for solving high throughput and reducing the pressure of the macrocell network. That is to say that SUs cannot influence the normal communication quality of MUs. In a non-cooperative way, users in the macrocell network

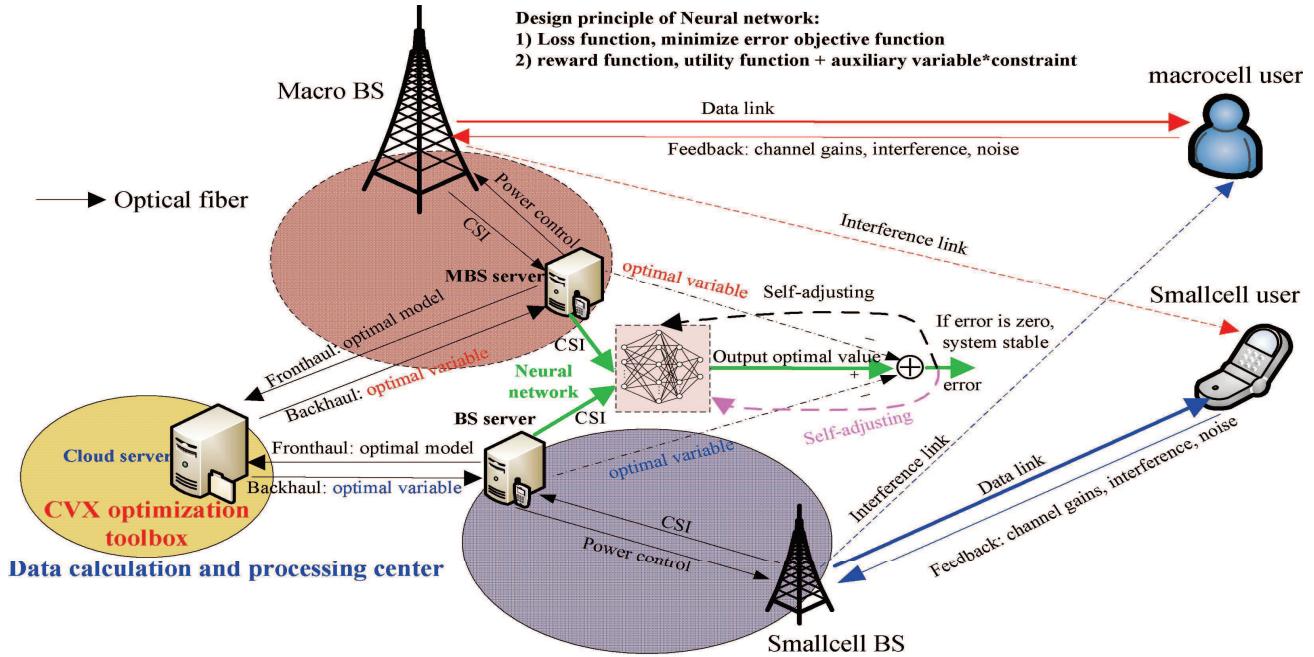


Fig. 13. A new learning-based structure of HetNets.

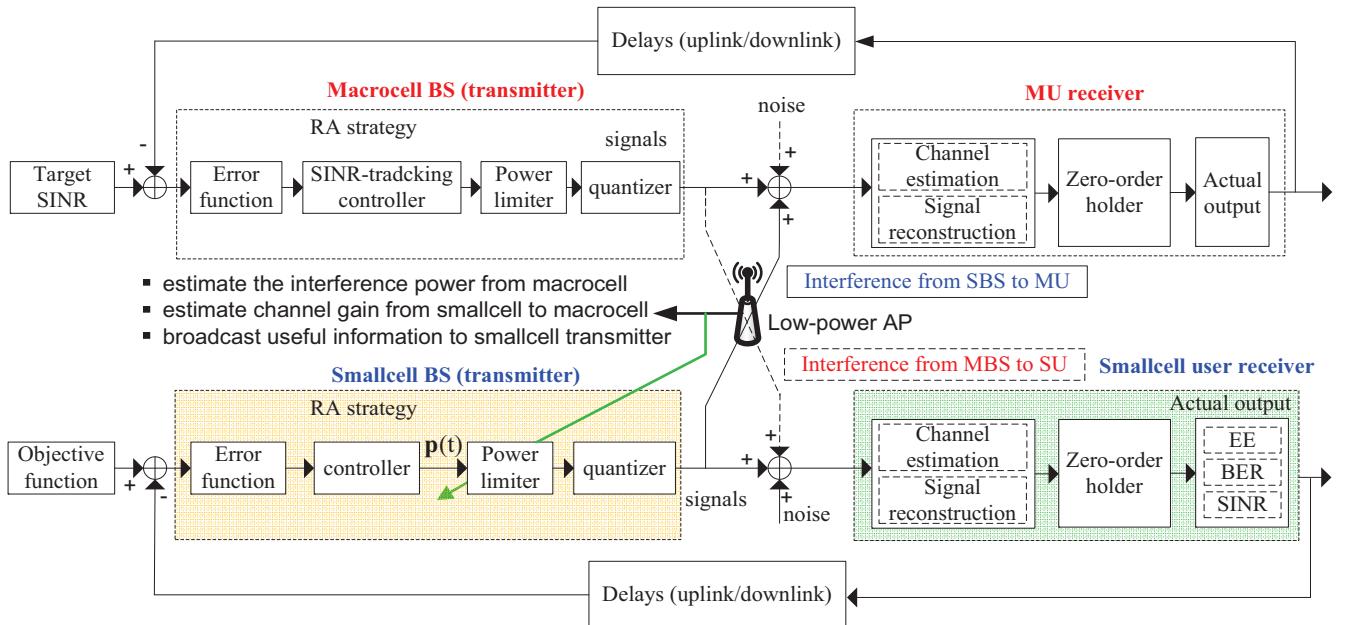


Fig. 14. A new control theory-based structure of HetNets.

have no obligation to provide any information for SUs. As an effective way, low-power APs can be used to estimate related messages for SUs' data transmission. The SBS can schedule all resources for obtaining a good performance improvement. The power-allocation controller can be designed and used to achieve optimization objectives as traditional optimization models.

(iii) *Low-power AP*: Due to the interference power constraint

of each MU, it is necessary to dynamically estimate the interference channel gain from SBS to MUs and feedbacks the information to the SBS, which is helpful to adjust the power limiter intelligently for protecting the QoS of MUs. Moreover, the AP estimates the MU's interference for SUs, so that the received SINR of SU's receiver can be easily calculated for further information feedback and performance optimization.

Comparing with traditional iteration-based RA schemes, control theory-based RA methods have many advantages which have been demonstrated in cognitive networks [221]–[226]. However, these methods can not directly extend to the HetNets due to the complexity of network structure. Additionally, due to the characteristics of control theory, the designed RA schemes can be well achieved by analog or digital circuits, e.g., proportional-integral-derivative (PID) controller. Therefore, control theory-based RAAs in HetNets have many benefits and may be the research direction of next-generation complicated HetNets.

G. Other Open Issues

With the mass of data requirements and the development of IoT, the whole wireless communication system has a great change, such as the V2V network or the M2M network. For different application scenarios, the RA and power optimization problems are different. The optimization problem may become a multiple-variable one. For example, in a HVN with MEC technology, our target is not only to reasonably adjust the transmit power of the user but also needs to consider the caching optimization and computation offloading in the communication system. In this case, we need to pay attention to the more practical and complicated application environment. Additionally, from the aspect of the solution process, the more intelligent and self-optimization algorithm should be introduced and designed for future HetNets, such as machine learning in a wireless communication application. These algorithms can adaptively match the surrounding radio scenes. We do not need to build an optimization model with multiple constraints and try to transform it into a convex form that can be efficiently solved by convex optimization theory or game theory. The training system can dynamically adjust its optimization parameters (e.g., transmit power, subcarrier assignment, beamforming matrix) to adapt the requirements of the wireless networks. With these intelligent algorithms (i.e., artificial-intelligent based schemes), the RA problems in HetNets will obtain a good solution in the future.

VI. CONCLUSION

In this article, a detailed survey on RA issues has been done for HetNets. The network structures and network scenarios of HetNets have been given together. The comparison of RA models in typical scenarios was presented. Then the state-of-art RAAs were introduced from theory (e.g., assume perfect CSI) to practical application (i.e., consider uncertainties or errors), including network scenarios, optimization objectives, approaches and so on. In addition, challenges and future trends were provided from optimization schemes to intelligent algorithms. It is expected that the RA in HetNets will play an important role in the system design of next-generation wireless communication for providing seamless connection, high system capacity and large-scale user access.

VII. APPENDIX

Table X lists the acronyms used in this survey

REFERENCES

- [1] A. Osseiran, et al., “Scenarios for 5G mobile and wireless communications: the vision of the METIS project,” *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.
- [2] C. Wang et al., “Cellular architecture and key technologies for 5G wireless communication networks,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.
- [3] M. De Ree, G. Mantas, A. Radwan, S. Mumtaz, J. Rodriguez and I. E. Otung, “Key management for beyond 5G mobile small cells: a survey,” *IEEE Access*, vol. 7, pp. 59200–59236, 2019.
- [4] A. Gupta and R. K. Jha, “A survey of 5G network: architecture and emerging technologies,” *IEEE Access*, vol. 3, pp. 1206–1232, 2015.
- [5] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios and J. Zhang, “Overview of millimeter wave communications for fifth-generation (5G) wireless networks with a focus on propagation models,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.
- [6] M. Agiwal, A. Roy and N. Saxena, “Next generation 5G wireless networks: a comprehensive survey,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, thirdquarter 2016.
- [7] F. Boccardi, R. W. Heath, A. Lozano, et al., “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [8] G. Karagiannis et al., “Vehicular networking: a survey and tutorial on requirements, architectures, challenges, standards and solutions,” *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 584–616, Fourth Quarter 2011.
- [9] A. Asadi, Q. Wang and V. Mancuso, “A survey on device-to-device communication in cellular networks,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, Fourthquarter 2014.
- [10] G. Wu, S. Talwar, K. Johnsson, N. Himayat, et al., “M2M: from mobile to embedded internet,” *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 36–43, Apr. 2011.
- [11] A. Al-Fuqaha, M. Guizani, M. Mohammadi, et al., “Internet of things: a survey on enabling technologies, protocols, and applications,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, Fourthquarter 2015.
- [12] A. Checko et al., “Cloud RAN for mobile networks—a technology overview,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, Firstquarter 2015.
- [13] P. Mach and Z. Becvar, “Mobile edge computing: a survey on architecture and computation offloading,” *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, Thirdquarter 2017.
- [14] K. Kumar and Y. Lu, “Cloud computing for mobile users: can offloading computation save energy?,” *Computer*, vol. 43, no. 4, pp. 51–56, April 2010.
- [15] Y. Zeng, R. Zhang and T. J. Lim, “Wireless communications with unmanned aerial vehicles: opportunities and challenges,” *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.
- [16] S. Abdelwahab, B. Hamdaoui, M. Guizani and A. Rayes, “Enabling smart cloud services through remote sensing: an internet of everything enabler,” *IEEE Internet of Things J.*, vol. 1, no. 3, pp. 276–288, Jun. 2014.
- [17] G. Wu, M. Mizuno and P. J. M. Havinga, “MIRAI architecture for heterogeneous network,” *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 126–134, Feb. 2002.
- [18] 3GPP TR 36.932 (V12.1.0), “Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN (Release 12),” 2013.
- [19] S. Parkvall, A. Furuskar and E. Dahlman, “Evolution of LTE toward IMT-advanced,” *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 84–91, Feb. 2011.
- [20] Y. Niu, Y. Li, D. Jin, et al., “A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges,” *Wireles Netw.*, vol. 21, no. 8, pp. 2657–2676, Nov. 2015.
- [21] Y. L. Lee, T. C. Chuah, J. Loo, et al., “Recent advances in radio resource management for heterogeneous LTE/LTE-A networks,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2142–2180, Fourthquarter 2014.
- [22] C. Barakat, E. Altman and W. Dabbous, “On TCP performance in a heterogeneous network: a survey,” *IEEE Commun. Mag.*, vol. 38, no. 1, pp. 40–46, Jan. 2000.
- [23] G. Karopoulos, G. Kambourakis and S. Gritzalis, “Survey of secure handoff optimization schemes for multimedia services over all-IP wireless heterogeneous networks,” *IEEE Commun. Surveys Tuts.*, vol. 9, no. 3, pp. 18–28, Third Quarter 2007.
- [24] R. Chai, W.-G. Zhou, Q.-B. Chen, et al., “A survey on vertical handoff decision for heterogeneous wireless networks,” in: *IEEE ICT*, Beijing, 2009, pp. 279–282.

- [194] H. Gacanin, "Autonomous wireless systems with artificial intelligence: A knowledge management perspective," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 51–59, 2019.
- [195] J. Sun, W. Shi, Z. Han, J. Yang, and G. Gui, "Behavioral modeling and linearization of wideband RF power amplifiers using BiLSTM networks for 5G wireless systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10348–10356, 2019.
- [196] H. Huang, S. Guo, G. Gui, Z. Yang, J. Zhang, H. Sari and F. Adachi, "Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions," *IEEE Wirel. Commun. Mag.*, in press, doi: 10.1109/MWC.2019.1900027.
- [197] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou, and D. Zhao, "Flight delay prediction based on aviation big data and machine learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1065–1069, 2020.
- [198] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, Sept. 2018.
- [199] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1065–1069, 2020.
- [200] N. Ye, X.-M. Li, H. Yu, L. Zhao, W. Liu, X. Hou, "DeepNOMA: A unified framework for NOMA using deep multi-task learning," *IEEE Trans. Wireless Commun.*, doi: 10.1109/TWC.2019.2963185
- [201] H. Liang, X. Zhang, J. Zhang, Q. Li, S. Zhou, and L. Zhao, "A novel adaptive resource allocation model based on SMDP and reinforcement learning algorithm in vehicular cloud system," *IEEE Trans. Veh. Technol.*, vol. 68, No. 10, pp. 10018–10029, Oct. 2019.
- [202] J. Sun, J. Wang, L. Guo, J. Yang and G. Gui, "Adaptive deep learning aided digital predistorter considering dynamic envelope," *IEEE Trans. Veh. Technol.*, in press, doi: 10.1109/TVT.2020.2974506
- [203] G. Gui, M. Liu, F. Tang, N. Kato, and F. Adachi, "6G: Opening new horizons for integration of comfort, security and intelligence," *IEEE Wireless Commun. Mag.*, in press, doi: 10.1109/MWC.001.1900516.
- [204] X. Nie, Y. Zhao, Z. Li, et al., "Dynamic TCP initial windows and congestion control schemes through reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1231–1247, Jun. 2019.
- [205] Z. Xu, J. Tang, C. Yin, et al., "Experience-driven congestion control: when multi-path TCP meets deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1325–1336, Jun. 2019.
- [206] K. Xiao, S. Mao and J. K. Tugnait, "TCP-drinc: smart congestion control based on deep reinforcement learning," *IEEE Access*, vol. 7, no. 1, pp. 11892–11904, 2019.
- [207] X. Huang, T. Yuan, G. Qiao, et al., "Deep reinforcement learning for multimedia traffic control in software defined networking," *IEEE Network*, vol. 32, no. 6, pp. 35–41, Nov. 2018.
- [208] N. Taherkhani and S. Pierre, "Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm," *IEEE Trans. Intell. Transport. Syst.*, vol. 17, no. 11, pp. 3275–3285, Nov. 2016.
- [209] H. Huang, J. Yang, H. Huang, et al., "Deep Learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sept. 2018.
- [210] J. Dai and H. C. So, "Sparse bayesian learning approach for outlier-resistant direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 744–756, Feb. 2018.
- [211] Z. Liu, C. Zhang and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Ant. Prop.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.
- [212] H. Wang, X. Wang, L. Wan, et al., "Robust sparse bayesian learning for off-grid DOA estimation with non-uniform noise," *IEEE Access*, vol. 6, no. 1, pp. 64688–64697, 2018.
- [213] P. Chen, Z. Cao, Z. Chen, et al., "Off-grid DoA estimation using sparse bayesian learning in MIMO radar with unknown mutual coupling," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 208–220, Jan. 2019.
- [214] H. Zamani, H. Zayyani and F. Marvasti, "An iterative dictionary learning-based algorithm for DOA estimation," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1784–1787, September 2016.
- [215] F. Wen, J. Wang, J. Shi, G. Gui, "Auxiliary vehicle positioning based on robust DOA estimation with unknown mutual coupling," *IEEE Internet Things J.*, in press, doi:10.1109/IOT.2020.2979771
- [216] B. Liu, G. Gui, S. Matsushita, and L. Xu, "Adaptive filtering algorithm for direction-of-arrival (DOA) estimation with small snapshots," *Digital Signal Processing*, vol. 94, pp. 84–95, Nov. 2019.
- [217] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming [Online]. Available: <http://cvxr.com/cvx> Dec. 2010, version 1.21.
- [218] A. Chockalingam, P. Dietrich, L. B. Milstein, et al., "Performance of closed-loop power control in DS-CDMA cellular systems," *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 774–789, Aug. 1998.
- [219] P.-R Chang and B.-C. Wang, "Adaptive fuzzy power control for CDMA mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 225–236, May 1996.
- [220] B.-K. Lee, Y.-H. Chen and B.-S. Chen, "Robust H_∞ power control for CDMA cellular communication systems," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3947–3956, Oct. 2006.
- [221] P. Setoodeh and S. Haykin, "Robust transmit power control for cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 915–939, May 2009.
- [222] G. Matsui, T. Tachibana, Y. Nakamura, et al., "Distributed power adjustment based on control theory for cognitive radio networks," *Computer Nets.*, vol. 57, no. 17, pp. 3344–3356, Dec. 2013.
- [223] M. Liu, J. Yang, T. Song, J. Hu, and G. Gui, "Deep learning-inspired message passing algorithm for efficient resource allocation in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 641–653, Jan. 2019.
- [224] S. Pan, X. Zhao and Y. Liang, "Robust power allocation for OFDM-based cognitive radio networks: A switched affine based control approach," *IEEE Access*, vol. 5, no. 1, pp. 18778–18792, 2017.
- [225] S. Zhang and X. Zhao, "Distributed power allocation based on LQG regulator with adaptive weight and switching scheme for cognitive radio networks," *IEEE Access*, vol. 6, no. 1, pp. 39180–39196, 2018.
- [226] S. Zhang and X. Zhao, "Distributed power allocation for cognitive radio networks with time varying channel and delay H-infinite state feedback control approach," *IEEE Access*, vol. 6, no. 1, pp. 56893–56910, 2018.