User Association Optimisation in HetNets: Algorithms and Performance

by

Dantong Liu

Doctor of Philosophy

School of Electronic Engineering and Computer Science
Queen Mary University of London
United Kingdom

November 2015

Abstract

The fifth generation (5G) mobile networks expect significantly higher transmission rate and energy efficiency than existing networks. Heterogeneous networks (HetNets), where various low power base stations (BSs) are underlaid in a macro-cellular network, are likely to become the dominate theme during the wireless evolution towards 5G. However the complex HetNets scenario poses substantial challenges to the user association design.

This thesis focuses on the user association optimisation for different HetNets scenarios.

First, user association policy is designed for conventional grid-powered HetNets via game theory. An optimal user association algorithm is proposed to improve the downlink (DL) system performance. In order to address the uplink-downlink (UL-DL) asymmetry issue in HetNets, a joint UL and DL user association algorithm is further developed to enhance both UL and DL energy efficiencies. In addition, an opportunistic user association algorithm in multi-service HetNets is proposed for quality of service (QoS) provision of delay constraint traffic while providing fair resource allocation for best effort traffic.

Second, driven by increasing environmental concerns, user association policy is designed for green HetNets with renewable energy powered BSs. In such a scenario, the proposed adaptive user association algorithm is able to adapt the user association decision to the amount of renewable energy harvested by BSs.

Third, HetNets with hybrid energy sources are investigated, as BSs powered by both power grid and renewable energy sources have the superiority in supporting uninterrupted service as well as achieving green communications. In this context, an optimal user association algorithm is developed to achieve the tradeoffs between average traffic delay and on-grid energy consumption. Additionally, a two-dimensional optimisation on user association and green energy allocation is proposed to minimise both total and peak

on-grid energy consumptions, as well as enhance the QoS provision.

Thorough theoretical analysis is conducted in the development of all proposed algorithms, and performance of proposed algorithms is evaluated via comprehensive simulations.

Acknowledgments

The most important person I would like to express my great gratitude to is my primary supervisor, Dr. Yue Chen. It is Dr. Yue Chen who introduced me to the research field and provided me with the invaluable guidance as well as the persistent encouragement, helping me go through all difficulties I encountered during my Ph.D. study. Her strongest support to both my research and daily life makes me not only grow to an independent and hard-working researcher, but also build up the well-meaning personality, just like what she has. I can never reach this level and stage without her. To be her student is the greatest treasure I have ever had.

I would like to express my appreciation to Dr. Kok Keong Chai, Dr. Tiankui Zhang, Dr. Maged Elkashlan and Prof. Laurie Cuthbert for their valuable suggestions and comments on my research, which are indispensable for my research progress.

I would also like to thank Prof. Qiang Ni (Lancaster University) and Prof. Mischa Dohler (King's College London) for agreeing to be my Ph.D. examiners.

My study would not have been complete without the help and the friendship of others, Dr. Lexi Xu, Dr. Fei Peng, Dr. Nan Wang, Dr. Yue Liu, Dr. Bo Zhong, Dr. Lifeng Wang, Dr. Yansha Deng, Xiuxian Lao, Xingyu Han, Dan Zhao, Yun Li, Xinyue Wang, Zhijin Qin, Liumeng Song, Aini Li, Xueke Lv, Yuanwei Liu, Jie Deng, Jingjing Zhao, Anqi He, Bingyu Xu, Yanru Wang, and Yuhang Dai among others. They always hold a place in my happy memories of my Ph.D. study in London.

Finally, with my love and gratitude, I would like to dedicate this thesis to my family who always give me the unreserved support and the selfishless love, especially to my parents and my boyfriend.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

5G	Fifth Generation
BA	Bandwidth Allocation
BS	Base Station
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CDF	Cumulative Distribution Function
CoMP	Coordinated Multipoint
C-RAN	Cloud Radio Access Network
CR	Cognitive Radio
CRE	Cell Range Extension
DL	Downlink
EA	Energy Allocation
EE	Energy Efficiency
eICIC	enhanced Inter-Cell Interference Coordination
FDD	Frequency-Division Duplex
H2H	Human-to-Human
HetNets	Heterogeneous Networks
ICIC	Inter-Cell Interference Coordination

IoT	Internet of Things
JFI	Jain's Fairness Index
JUDUA	Joint UL and DL User Association
KKT	Karush-Kuhn-Tucker
LTE-A	LTE-Advanced
M2M	Machine-to-Machine
MBS	Macro Base Station
MIMO	Multiple-Input Mutiple-Output
mmWave	Millimeter Wave
NBS	Nash Bargaining Solution
OFDMA	Orthogonal-Frequency-Division-Multiple-Access
OPEX	Operational Expenditure
PC	Power Control
PDF	Probability Density Function
PBS	Pico Base Station
QoS	Quality of Service
RSS	Received Signal Strength
SD	Standard Deviation
SE	Spectrum Efficiency
SINR	Signal-to-Interference-plus-Noise Ratio
SP	Spectrum Partitioning
SNR	Signal-to-Noise Ratio
TDD	Time-Division Duplex
UA	User Association
UE	User Equipment
UL	Uplink

Chapter 1

Introduction

1.1 Background

The sky-rocketing proliferation of multimedia infotainment applications and high-end devices (e.g., smartphones, tablets, wearable devices, laptops, machine-to-machine communication devices) exacerbates the stringent demand for high data rate services. According to the latest visual network index (VNI) report from Cisco [Cis15], the global mobile data traffic will increase nearly tenfold between 2014 and 2019, reaching 24.3 exabytes per month by 2019 as shown in Fig. 1.1, wherein nearly three-fourths will be video, indicated by Fig. 1.2. The researchers in the field of communications have reached a consensus that incremental improvements to today's communication systems cannot meet the data demands in the foreseeable future. A paradigm shift is required for the emerging fifth generation (5G) mobile networks [ABC+14].

The goals of 5G are broad, but are presumed to include the provision of at least 1,000 times higher wireless area capacity than current mobile networks. Other high-level key performance indicators (KPIs) envisioned by 5G-Public Private Partnership (5G-PPP) [5gp] include 10 times lower energy consumption per service, reduction of the average service creation time cycle from 90 hours to 90 minutes, creation of a secure, reliable and dependable Internet with a "zero perceived" downtime for service provision, facilitation of very dense deployment of wireless communication links to connect

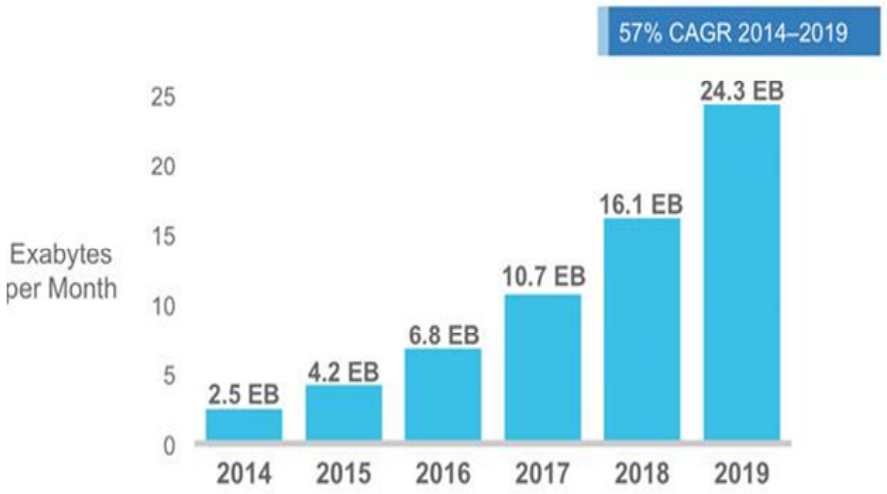


Figure 1.1: Cisco forecasts 24.3 exabytes per month of mobile data traffic by 2019 [Cis15].

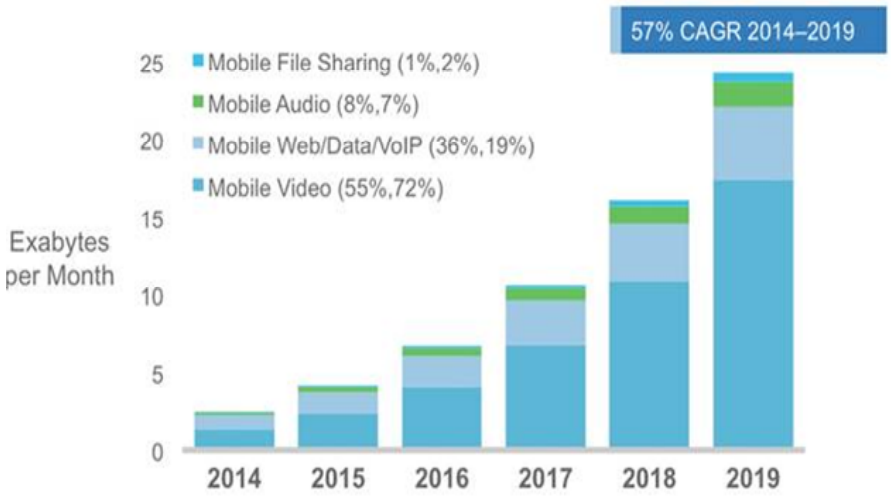


Figure 1.2: Mobile video will generate about 72 percent of mobile data traffic by 2019 [Cis15].

over 7 trillion wireless devices serving over 7 billion people, and enabling advanced user controlled privacy [5gp]. To address these, the primary technologies and approaches identified by E. Hossain *et al.* [HH15] for 5G networks are dense heterogeneous networks (HetNets), device-to-device communication, full-duplex communication, massive multiple-input multiple-output massive (MIMO) and millimeter wave (mmWave) communications technologies, energy-aware communication and energy harvesting, cloud-based radio access network (C-RAN) and visualisation of wireless resources. More specif-

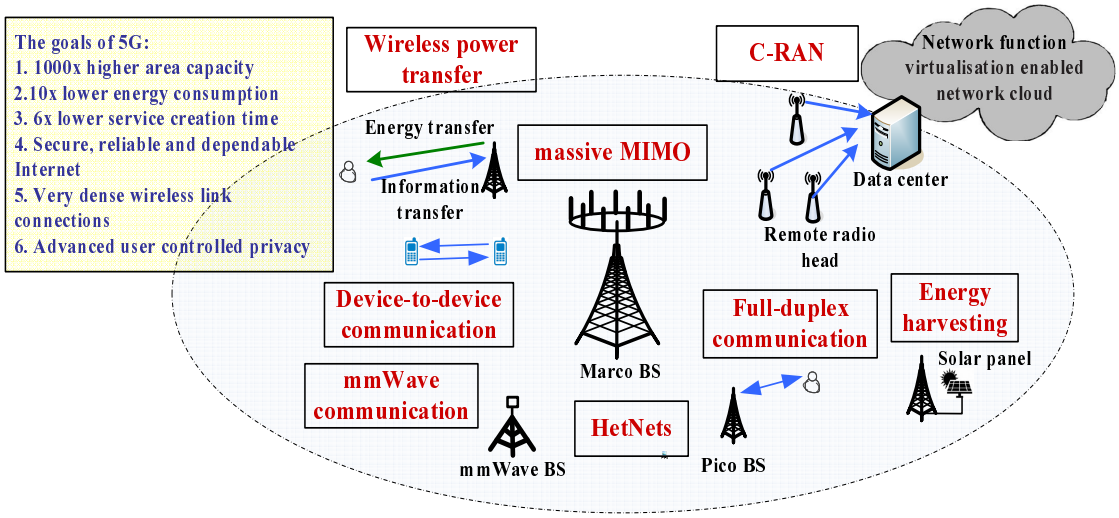


Figure 1.3: Enabling technologies and expected goals for 5G networks.

ically, J. G. Andrews *et al.* spotlight HetNets, mmWave, and massive MIMO as the "big three" 5G technologies [ABC+14]. Fig. 1.3 illustrates the enabling technologies and expected goals for 5G networks.

User association, namely associating a user with a particular serving base station (BS), substantially affects the network performance. In the mature LTE/LTE-Advanced (LTE-A) systems, the radio admission control entity is located at the radio resource control entity in Layer 3 of the protocol stack, which decides whether a new radio-bearer admission request should be admitted or rejected. The decision is made according to the quality of service (QoS) requirements of the requesting radio bearer, the priority level of the request and the availability of radio resources, with the aim to maximise the radio resource utilisation [3gp15]. In existing systems, the received power based user association rule is widely used [DOC10], where a user will choose to associate with the specific BS, from which it obtains the maximum received signal strength (max RSS) in downlink (DL) direction. Nevertheless, the inherent nature of the 5G technologies inevitably makes such a rudimentary user association rule ineffective. Hence, user association algorithms should be redesigned for 5G networks.

In this thesis, a great emphasis is given to the user association design in HetNets.

1.2 Research Contributions

The contributions of the thesis are summarised as follows.

- An extensive and detailed overview of the state-of-the-art user association in grid-powered HetNets and renewable energy powered networks is carried out. Additionally, open challenges of user association in this context are highlighted, which sheds lights on the research direction.

- In conventional grid-powered HetNets, a Nash Bargaining Solution (NBS) based user association algorithm is first proposed to improve the network DL performance. Then in order to address the uplink-downlink (UL-DL) asymmetry in HetNets, a joint UL and DL user association algorithm using NBS is developed to enhance both UL and DL energy efficiencies. Subsequently, taking multi-service into consideration, a NBS based opportunistic user association algorithm is proposed for QoS provision of delay constraint traffic while providing fair resource allocation for best effort traffic.

- In green HetNets with renewable energy powered BSs, both optimal offline and heuristic online algorithms are designed for the adaptive user association, which are able to adjust the user association decision according to the amount of renewable energy harvested by BSs.

- Finally, the emerging paradigm with hybrid energy sources is investigated, where BSs in HetNets are powered by both power grid and renewable energy sources. An optimal user association algorithm is developed to achieve the tradeoffs between average traffic delay and on-grid energy consumption. In addition, two-dimensional optimisation on user association and green energy allocation is carried out to minimise both total and peak on-grid energy consumptions, as well as enhance the QoS provision.

- Thorough theoretical analysis is conducted in the development of all proposed

algorithms, and the effectiveness of all proposed algorithms is validated via comprehensive simulations.

1.3 Author's Publication

Journal papers

1. **Dantong Liu**, Lifeng Wang, Yue Chen, Maged Elkashlan, Kai-Kit Wong, Robert Schober, and Lajos Hanzo, "User Association in 5G Networks: A Survey and an Outlook," *Submitted to IEEE Communications Surveys & Tutorials* (under minor revision), Nov. 2015.

2. **Dantong Liu**, Yue Chen, Kok Keong Chai, Tiankui Zhang and Maged Elkashlan, "Two-Dimensional Optimization on User Association and Green Energy Allocation for HetNets with Hybrid Energy Sources," *IEEE Transactions on Communications*, vol.63, no.11, pp.4111-4124, Nov. 2015.

3. **Dantong Liu**, Lifeng Wang, Yue Chen, Tiankui Zhang, Kok Keong Chai and Maged Elkashlan, "Distributed Energy Efficient Fair User Association in Massive MIMO Enabled HetNets", *IEEE Communications Letters*, vol.19, no.10, pp.1770-1773, Oct. 2015.

4. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Backhaul Aware Joint Uplink and Downlink User Association for Delay-Power Tradeoffs in HetNets with Hybrid Energy Sources," *Transactions on Emerging Telecommunications Technologies*, DOI: 10.1002/ett.2968, Jul. 2015.

5. **Dantong Liu**, Yue Chen, Kok Keong Chai, Tiankui Zhang and Maged Elkashlan, "Opportunistic User Association for Multi-Service HetNets Using Nash Bargaining Solution," *IEEE Communications Letters*, Volume 18, Issue 3, Page(s):463-466, Mar. 2014.

6. **Dantong Liu**, Yue Chen, Tiankui Zhang, Kok Keong Chai, Loo, J. and Vinel, A., "Stackelberg Game Based Cooperative User Relay Assisted Load Balancing in Cellular Networks," *IEEE Communications Letters*, Volume 17, Issue 2, Page(s):424-427, Feb. 2013.

Conference papers

1. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Joint User Association and Green Energy Allocation in HetNets with Hybrid Energy Sources," *IEEE Wireless Communications and Networking Conference (WCNC 2015)*, New Orleans, LA, USA, Mar. 2015.

2. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Distributed Delay-Energy Aware User Association in 3-tier HetNets with Hybrid Energy Sources," *Globecom 2014 - 2nd Workshop on Green Broadband Access: Energy Efficient Wireless and Wired Network Solutions*, Austin, Texas, USA, Dec. 2014.

3. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Optimal User Association for Delay-Power Tradeoffs in HetNets with Hybrid Energy Sources," *International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2014)*, Washington DC, USA, Sept. 2014.

4. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Joint Uplink and Downlink User Association for Energy-Efficient HetNets Using Nash Bargaining Solution," *2014 IEEE 79th Vehicular Technology Conference (VTC2014-Spring)*, Seoul, Korea, May 2014.

5. **Dantong Liu**, Yue Chen, Kok Keong Chai, Tiankui Zhang and Chengkang Pan, "Adaptive User Association in HetNets with Renewable Energy Powered Base Stations," *International Conference on Telecommunications 2014 (ICT 2014)*, Lisbon, Portugal, May 2014.

6. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Nash Bargaining Solution Based User Association Optimization in HetNets," *11th Annual IEEE Consumer Communications Networking Conference (CCNC 2014)*, Les Vegas, Nevada, USA, Jan. 2014.

7. **Dantong Liu**, Yue Chen, Kok Keong Chai and Tiankui Zhang, "Performance evaluation of Nash Bargaining Based User Association in HetNet," *9th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2013)*, Lyon, France, Oct. 2013.

8. **Dantong Liu**, Yue Chen, and Kok Keong Chai, "Cooperative User Relaying Assisted Load Balancing Scheme for OFDMA Based Cellular Networks," *2012 3rd IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)* , Beijing, China, Sept. 2012.

9. **Dantong Liu**, Yue Chen, Kok Keong Chai, Laurie Cuthbert, and Tiankui Zhang, "Cognitive Cooperative Traffic Offloading Scheme over Heterogeneous Networks," *2012 7th International ICST Conference on Communications and Networking in China (CHINACOM)*, Kunming, China, Aug. 2012.

10. **Dantong Liu**, Yue Chen, and Tiankui Zhang, "Research on Cooperative User Relaying Assisted Load Balance Scheme in OFDMA Networks," *2012 International Workshop on Information and Electronics Engineering*, Harbin, China, Jan. 2012.

1.4 Thesis Organisation

Chapter 2 introduces fundamental concepts of HetNets, summarises the state-of-the-art user association in conventional grid-powered HetNets and renewable energy powered networks, as well as highlights open challenges of user association in this context.

Chapter 3 proposes three algorithms in conventional grid-powered HetNets, includ-

ing NBS based user association for DL performance optimisation, joint UL and DL user association using NBS, and opportunistic user association in multi-service HetNets. The theoretical analysis and performance evaluation are carried out for all these three proposed algorithms.

Chapter 4 presents the adaptive user association in HetNets with renewable energy powered BSs, where both optimal offline and heuristic online algorithms are developed. The theoretical analysis and performance evaluation of proposed algorithms provide guidelines on the user association policy design in the scenario where all BSs are solely powered by renewable energy sources.

Chapter 5 investigates the emerging paradigm with hybrid energy sources, where BSs in HetNets are powered by both power grid and renewable energy sources. It begins with the proposed optimal user association algorithm to achieve the tradeoffs between average traffic delay and on-grid energy consumption. Afterwards, it presents the work about two-dimensional optimisation on user association and green energy allocation to minimise both total and peak on-grid energy consumptions, as well as enhance the QoS provision. The theoretical analysis and performance evaluation are conducted during the development of proposed algorithms.

Chapter 6 consists of conclusions and future work.

Chapter 2

Fundamental Concepts and State-of-the-Art

2.1 Overview of HetNets

Wireless data traffic increases exponentially in recent years due to the proliferation of advanced user terminals and bandwidth-greedy applications (e.g., smart phones, tablets, mobile TV). In order to enhance the performance of the overall network, LTE-A standardisation proposed the use of advanced technologies [MDO09], such as carrier aggregation (CA), MIMO and coordinated multipoint (CoMP). CA allows the concurrent utilisation of different frequency carriers, hence efficiently increasing the bandwidth that can be allocated to end users. The enhancement of multi-antenna techniques, where using MIMO systems with up to 8*8 antenna arrays has been regarded as an effective way to improve the system capacity. CoMP transmission and reception, where multiple cells are able to coordinate their scheduling or transmission to serve users with adverse channel conditions, is also envisioned to notably mitigate outage probability at the cell edge. However, all these advanced technologies cannot bring in significant enhancements as they are reaching theoretical limits in conventional homogeneous networks [LPGDlR+11].

Further improvement in system spectrum efficiency seems only possible by increas-

ing the node deployment density–cell splitting/densification [DMW+11]. In a relatively sparse deployment of macro base stations (MBSs), adding another BS does not severely increase the intercell interference, and solid cell splitting gains can be easily achieved. However, in today's dense deployments, cell splitting gains are significantly reduced due to the already severe intercell interference. Moreover, site acquisition costs in a capacity limited dense urban area can be prohibitively expensive [DMW+11].

Issues associated with the above traditional macrocell splitting/densification can be overcome by multi-tier HetNets roll-outs, which is envisaged by LTE standardisation. HetNets involve small cells underlaying the existing macro cellular networks. Small cells, such as picocells, femtocells and relays, transmit with less power and serve as fundamental elements for the traffic offloading from macrocells, thereby improving the network coverage, enhancing the cell-edge user performance and boosting both the spectrum and energy efficiencies [LPGDlR+11]. On the other hand, such a new palette of low power "miniature" BSs in small cells requires less upfront planning and lease costs, and consequently substantially reduces the network operational and capital expenditures (OPEX, CAPEX) [CAG08]. The following details the features of small cells in HetNets.

- Picocells are covered by low-power operator-installed BSs relying on the same backhaul and access features as macrocells. They are usually deployed in a centralized manner, serving a few tens of users within a radio range of 300 m or less, and have a typical transmit power range from 23 to 30 dBm [LPGDlR+11]. Picocells do not require an air conditioning unit for the power amplifier, and incur much lower cost than traditional MBSs [DMW+11].

- Femtocells, also known as home BSs or home eNBs, are low-cost, low-power and user-deployed access points. Typically, a femtocell range is less than 50 m and its transmit power is less than 23 dBm. They operate in open or restricted (closed subscriber group) access [LPGDlR+11].

- Relays are usually operator-deployed access points that route data from the MBS

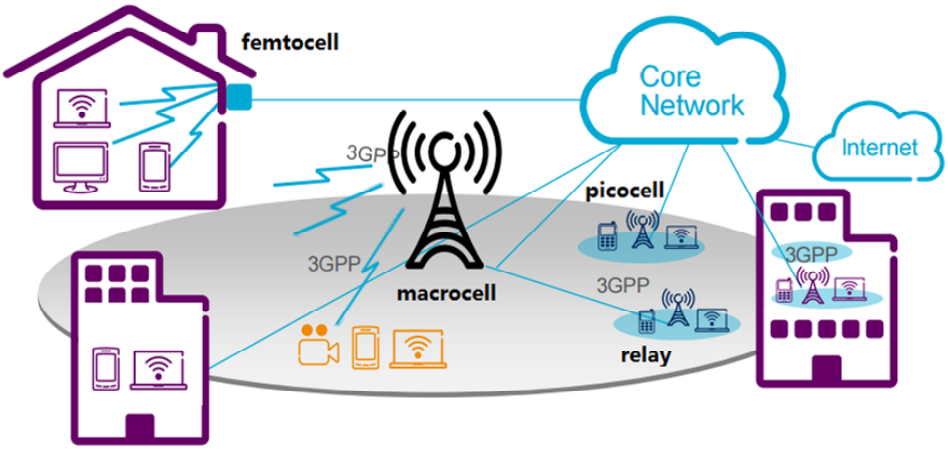


Figure 2.1: Architecture of HetNets [Eri12].

to users and vice versa [LPGDlR+11]. Relays connect to the rest of the network via wireless backhaul. They can be deployed indoors or outdoors, with the transmit power ranging from 23 to 33 dBm for outdoor deployment, and 20 dBm or less for indoor deployment [DMW+11].

Fig. 2.1 illustrates the architecture of HetNets. HetNets provide an evolutionary paradigm from the single-tier to the multi-tier network infrastructure.

Apart from the transmit power disparity between small cells and macrocells in HetNets, the inherent nature of HetNets also manifests itself in terms of the UL-DL asymmetry, the backhaul bottleneck, diverse footprints and so on.

- **UL-DL asymmetry**: HetNets introduce a major asymmetry between DL and UL. The optimal resource allocation algorithm for DL or UL only will be less effective for the opposite direction. Specifically, the max DL RSS based user association rule may associate a user with the faraway marcocell, rather than with the nearby small cell. As a result, the user has to transmit at a potentially excessive power for guaranteeing the target received signal strength in the UL, thereby inflicting a high uplink interference on the small cell users.

- **Backhaul bottleneck**: In HetNets the potentially densely deployed small BSs

may impose an overwhelming backhaul traffic. On the other hand, the current small cell backhaul solutions, such as xDSL and non-line-of-sight (NLOS) microwave, are far from an ideal backhaul solution owing to their limited data rate [NNB+13]. Hence, the backhaul data rate constraint has become far more important in HetNets.

- **Mobility support**: The reduced transmit powers of small cells lead to reduced footprints. As a result, for a user having moderate or high mobility, an algorithm that does not consider the mobility issues may result in more frequent handovers among the cells in HetNets compared to conventional homogeneous cellular networks.

All the inherent nature of HetNets imposes substantial challenges to the resource allocation, mobility management, network deployment, etc. This thesis focuses the attention on the user association optimisation in different HetNets scenarios.

2.2 User Association in Grid-Powered HetNets

HetNets have been explored as the dominant theme for wireless evolution into 5G [BLM+14]. In the conventional grid-powered HetNets, all BSs are powered by constant energy supply. The classic max RSS user association rule is not particularly suitable for HetNets, as the transmit power disparity of marcocells and small cells will lead to the association of most of the users with the MBS [DOC10], hence potentially resulting in inefficient small cell deployment.

To cope with this problem, the concept of biased user association was proposed by 3GPP in Release 10 [Kyo10], which is also known as the cell range extension (CRE) scheme. For biased user association, the users' received power from small BSs is artificially increased by adding a bias to ensure that more users will be associated with small cells. In [Guv11], off-loading benefits of biased user association were demonstrated in

terms of the capacity improvement. However, the drawback of biased user association is that the group of users, who are forced to be associated with small cells by the added bias, experience strong interference from the nearby macrocell [JSXA11]. In this context, the improvement achieved by offloading traffic to small cells might be offset by such strong interference. Therefore, the tradeoff between network load balancing and network throughput strictly depends on the value of the selected bias, which has to be carefully optimised in order to maximise the network utility [HRTA14]. Moreover, several interference mitigation schemes based on resource partitioning have been proposed to solve the above problem in biased user association, including inter-cell interference coordination (ICIC) proposed in the 3GPP Release 8, and enhanced inter-cell interference coordination (eICIC) proposed in the 3GPP Release 10 [Std10]. The authors of [DMMS14] optimised both the bias value and the resource partitioning in eICIC enabled HetNets.

In this section, the existing research results on user association in grid-powered HetNets are surveyed and categorised according to different performance metrics, including spectrum efficiency and energy efficiency. Additionally, all user association algorithms mentioned in this section are summarised in Table 2-A, which provides the qualitative comparison of the existing user association algorithms conceived for grid-powered HetNets. In Table 2-A, "-" means that the corresponding algorithm did not consider this metric, "UA" stands for user association, "SE" and "EE" are spectrum efficiency and energy efficiency, respectively, "PC" and "SP" represent power control and spectrum partitioning, respectively.

2.2.1 User Association for Spectrum Efficiency Optimisation

Spectrum efficiency refers to the information rate that can be transmitted over a given bandwidth in a specific communication system. With the surge of data traffic and limited spectrum resources, a high spectrum efficiency is a mandatory requirement of 5G networks [HQ14].

Table 2-A: Qualitative Comparison of User Association Algorithms for Grid-Powered HetNets

Ref.	Algorithm	Direction	Control	SE	EE	QoS	Fairness
[CFM12]	UA	DL	Centralised	High	-	-	Low
[YRC+13]	UA	DL	Distributed	Moderate	-	-	High
[MAAV14a]	UA	DL	Distributed	High	High	High	-
[XHWW14]	UA	DL	Centralised	High	High	High	-
[SSV+14]	UA	DL	Distributed	High	-	Moderate	High
[NSMV14, PBS+13]	UA	DL	Distributed	High	-	High	-
[SHZ+14]	UA	UL	Distributed	High	-	High	-
[FADR11, GR13]	UA+SP	DL	Centralised	Moderate	-	-	High
[SY14]	UA+PC	DL	Distributed	High	-	-	High
[MBS+10, SHL15]	UA+PC	DL	Distributed	High	-	High	High
[PMN13]	UA+PC	UL	Centralised	-	High	High	-
[ZWC12]	UA+PC	DL	Hybrid	-	High	High	-
[HL14]	UA+PC	UL	Distributed	High	-	High	High
[HL13]	UA+PC	UL	Distributed	High	-	-	High
[SYXM13]	UA+sleeping	DL	Centralised	-	High	Moderate	High
[CRAF15]	UA+sleeping	DL	Centralised	-	High	High	-
[SQKS13]	biased UA+ sleeping	DL	Distributed	-	High	High	-

In [CFM12], a dynamic user association algorithm was proposed for DL multi-cell HetNets to maximise the sum rate of all users. The authors derived an upper bound of the DL sum rate using convex optimisation, and then proposed a heuristic user association rule having a low complexity and approaching the upper bound. Simulation results verified the superiority of the proposed heuristic user association rule over classic max RSS and biased user association in terms of the average user data rate. However, it is well known that maximising the sum data rate of all users may result in an unfair data rate allocation, which is also reflected by the results in [CFM12, Fig. 3]. [CFM12, Fig. 3] shows that the load of small cells is much larger than that of macrocells, resulting in small cells that are congested. Consequently, only the users in the macrocell centre achieve high data rates, and the other users are starved. To cope with this, in [YRC+13], a low-complexity distributed user association algorithm was proposed for the sum logarithmic user data rate maximisation. Since the logarithm is concave, and has diminishing returns, allocating more resources to an already well-served user has lower priority, whereas providing more resources to users with lower rates is desirable, thereby encouraging load balancing and user fairness. In [YRC+13], by relaxing the primal deterministic user association to a fractional association, the intractable primal combinatorial optimisation was converted into a convex optimisation. By exploiting the convexity of the problem, a distributed user association algorithm was developed with

the assistance of dual decomposition and gradient descent method, which converged to the optimum solution under the guarantee of not exceeding a certain maximum discrepancy from optimality. It is noted that the convergence speed of the gradient descent method heavily depends on the particular choice of step size. For the same problem formulation as in [YRC+13], a coordinate descent method was proposed in [SY14] for providing a rigourous performance guarantee and faster convergence compared to the algorithm of [YRC+13].

Game theory is also widely applied in the context of user association for spectrum efficiency optimisation. [SSV+14, NSMV14, PBS+13] formulated the DL user association in HetNets as a many-to-one matching game, where users and BSs evaluated each other based on well defined utilities. In [SSV+14], users and BSs ranked one another based on utility functions that accounted for both the data rate and the fairness to cell edge users, which were captured by newly proposed priorities. Different from [SSV+14], which differentiates the user priority, in [NSMV14] the delivery time, handover failure probability and various QoS requirements of users were taken into consideration when designing the utility for the user association. With problem formulation similar to the one in [NSMV14], the authors of [PBS+13] specifically focused on multimedia data services and characterised the user's quality of experience in terms of mean opinion scores that accurately reflected characteristics of the wireless transmission and the data application. Another interesting work was [SHZ+14], where the UL user association in HetNets was formulated as a college admission game with transfers, where a number of colleges, i.e., BSs in macrocells and small cells, sought to recruit a number of students, i.e., users. The formulated college admission game well captures the users's need to optimise their packet success rates and delays, as well as the small cell's incentive to offload traffic from the macrocell and extend its coverage.

The densely deployed small cells exacerbate the stringent demand for interference management in HetNets. The joint optimisation of user association and other aspects of radio resource allocation understandably has prompted significant research efforts.

In [FADR11], joint optimisation of user association and channel allocation between macrocells and small cells was investigated to maximise the minimum user data rate. Extending from [FADR11], in [GR13] joint user association, transmission coordination and channel allocation between macrocells and small cells was proposed to maximise the user data rate related utility. Joint user association and power control was investigated for DL HetNets in [SY14, MBS$^+$10, SHL15] and UL HetNets in [HL14, HL13]. The algorithms proposed in [SY14, MBS$^+$10, SHL15, HL14] iteratively updated the user association solution and transmit power until convergence, while [HL13] formulated the sum throughput maximisation problem as a non-cooperative game, with both users and BSs acting as players. It is noticed that the aforementioned joint optimisation of user association and channel allocation/power control turns out to be NP-hard, hence finding the optimal solution is not trivial. The solution is approached by either updating the user association and the power level sequentially in an iterative manner until convergence as in [SY14, MBS$^+$10, SHL15, HL14], or solving the user association optimisation with the fixed channel allocation/transmission coordination/power control, and then optimising channel allocation/transmission coordination/power control accordingly, and vice versa, as in [FADR11, GR13, HL13]. As a result, it comes to the conclusion that careful user association optimisation is crucial for the holistic optimisation of HetNets, indisputably underlining the significance of rigourous investigations on user association optimisation.

The qualitative comparison of the above-mentioned user association algorithms for spectrum efficiency optimisation in grid-powered HetNets is detailed in Table 2-A.

2.2.2 User Association for Energy Efficiency Optimisation

The skyrocketing increase of data traffic volume and the dramatic expansion of network infrastructures will inevitably trigger a tremendous escalation of energy consumption in wireless networks, which will directly result in an increase of greenhouse gas emission and mandate an ever increasing urgency for protecting the environment and the sustainable network development. Consequently, both industry and academia are engaged in working

towards enhancing the network energy efficiency.

Maximising network energy efficiency may be defined as maximising the amount of successfully sent data while minimising the total energy consumption. As far as the problem formulation is concerned, maximising network energy efficiency can be either expressed as minimising the total energy consumption while satisfying user traffic demands or maximising the ratio between the total data rate of all users and the total energy consumption, which is defined as the energy efficiency (bits/Joule). Note that macrocells have a much higher transmit power than small cells, and thus the access network energy consumption is higher when a user is associated with a macrocell. Hence, the network energy efficiency is highly dependent on the user association decision [MAAV14b].

Numerous valuable contributions have been published on energy efficient user association in HetNets recently. In [PMN13], a user association algorithm was developed for UL HetNets in order to maximise the system energy efficiency subject to user's maximum transmit power and minimum rate constraints. In [MAAV14a], user association in DL HetNets was optimised for maximisation of the ratio between the total data rate of all users and the total energy consumption. For a problem formulation different with the one in [MAAV14a], in [ZWC12] the authors investigated energy efficient user association aiming at minimising the total power consumption while satisfying the users' traffic demand. The authors of [XHWW14] considered the user association for users with video applications, where a video content aware energy efficient user association algorithm was proposed for DL HetNets, with the aim to maximise the ratio of the peak-signal-to-noise ratio and the system energy consumption, and nonlinear fractional programming and dual decomposition techniques were adopted to resolve the problem.

Statistical studies of mobile communication systems have shown that 57% of the total energy consumption of wireless network operation can be attributed to the radio access nodes [CZXL11], where about 60% of the power at each BS is consumed by processing circuits and air conditioning [AGD+11]. As a result, shutting down BSs without active

users is believed to be an efficient way to reduce network power consumption [FJL+13, RF14].

In [SYXM13], joint optimisation of long-term BS sleeping, user association, and subcarrier allocation was considered for maximisation of the energy efficiency or minimisation of the total power consumption under constraints on the average sum rate and the user rate fairness. The performance of these two formulations (i.e., energy efficiency maximisation and total power minimisation) was investigated via simulations. In [CRAF15] an energy efficient algorithm was introduced for minimisation of the energy consumption by adjusting user association and BSs sleeping policies, where the dependence of the energy consumption on the spatio-temporal variations of traffic demands and internal hardware components of BSs were fully considered. In addition, in [SQKS13], the coverage probability and the energy efficiency in K-tier heterogeneous wireless networks were derived under different sleeping policies using a stochastic geometry based model, where the authors formulated power consumption minimisation and energy efficiency maximisation problems, and determined the optimal operating regimes of the macrocell.

The qualitative comparison of the above-mentioned user association algorithms for energy efficiency optimisation in grid-powered HetNets is detailed in Table 2-A.

2.3 User Association in Renewable Energy Powered Networks

One of the main challenges in 5G networks is the improvement of the energy efficiency of radio access networks (RANs). Recently, motivated by environmental concerns and the regulatory pressure for greener solutions, network operators have elaborated on considering the deployment of renewable energy sources, such as solar panels and wind turbines, to supplement the conventional power grid in powering BSs. In this scenario, BSs are capable of harvesting energy from the environment, and do not require an *always-on*

Table 2-B: Qualitative Comparison of User Association Algorithms for Renewable Energy Powered Networks.

Ref.	Algorithm	Direction	Control	SE	EE	QoS	Fairness
[DLN+14]	max-RSS UA	DL	Distributed	-	High	-	-
[SZZH14]	biased UA	DL	Distributed	High	High	-	-
[RPIdOV14]	UA	DL	Distributed	Moderate	High	High	High
[HA12]	UA	DL	Centralized	-	High	High	-
[HA13]	UA+EA	DL	Centralized	-	High	High	-
[WKLY15]	UA+BA	DL	Distributed/ Centralized	-	High	High	-
[JZLL15]	UA+PC+BA	UL	Centralized	High	High	High	-

energy source [HN13]. This is of considerable interest in undeveloped areas where the power grid is not readily available, as well as to open up entirely new categories of low cost *drop and play* small cell deployments rather than the *plug and play*.

In this section, the existing research results on user association in renewable energy powered networks are surveyed and categorised according to different network scenarios, including solely renewable energy powered networks and hybrid energy powered networks. In addition, all user association algorithms mentioned in this section are summarised in Table 2-B, which provides the qualitative comparison of the existing user association algorithms for renewable energy powered networks, where "UA" stands for user association, "SE" and "EE" are spectrum efficiency and energy efficiency, respectively, "PC" is power control, "BA" and "EA" represent bandwidth allocation and energy allocation, respectively.

2.3.1 User Association in Solely Renewable Energy Powered Networks

Due to the randomness of the energy availability in renewable energy sources, integrating energy harvesting capabilities in BSs entails many challenges for user association algorithm design. The user association decision should be adapted according to the energy and load variations across time and space. The authors of [DLN+14] developed a traceable model for HetNets via stochastic geometry where all BSs were assumed solely powered by renewable energy sources, and provided a fundamental characterisation of regimes under which HetNets with renewable energy powered BSs have the same per-

formance as the ones with grid-powered BSs. By relaxing the primal deterministic user association to a fractional user association, the authors of [RPIdOV14] proposed a user association algorithm for maximisation of the DL aggregate network utility based on per-user throughput, where BSs were solely powered by renewable energy and equipped with finite capacity batteries.

The deployment of relays with energy harvesting capabilities has also attracted significant attention recently, as it is able to improve the system capacity and the coverage in remote areas without ready access to the power grid. In this context, [SZZH14] studied the user association problem targeting DL throughput optimisation for energy harvesting relay-assisted cellular networks, where BSs were powered by the power grid, and relays were powered by the renewable energy. Using tools from stochastic geometry, the authors developed a dynamic biased user association algorithm where the bias was optimised based on renewable energy arrival rates. In [JZLL15], joint user association, resource block allocation and power control was investigated for maximisation of the UL network throughput in cellular networks with energy harvesting relays. The energy-harvesting process was stochastically described by a Poisson process. The authors proposed a new metric termed as the survival probability as the selection criterion for an energy-harvesting relay to support data transmissions.

The qualitative comparison of the above-mentioned user association algorithms for solely renewable energy powered networks is detailed in Table 2-B.

2.3.2 User Association in Hybrid Energy Powered Networks

Although the amount of renewable energy is potentially unlimited, the intermittent nature of the energy from renewable energy sources results in a highly random energy availability at the BS. Thus BSs powered by hybrid energy sources, which employ a combination of power grid and renewable energy sources, are preferable over those solely powered by renewable energy sources in order to support uninterrupted service [NLS13]. The

concept of hybrid energy sources has already been adopted by the industry. For instance, Huawei has deployed BSs which draw their energy from both constant energy supplies and renewable energy sources [HUA]. In the literature, power allocation [GZN13], coordinated MIMO [CLW12] and network planning [ZPSY13] have been studied in the context of cellular networks powered by hybrid energy sources. For the user association designed for such networks, the vital issue is to minimise the on-grid energy consumption as well as guarantee the user QoS provision, see [HA12, WKLY15, HA13] and references therein. In [HA12], an intelligent cell breathing method was proposed to minimise the maximal green energy depletion rate of BSs. In [WKLY15], a constrained total energy cost minimisation problem was formulated, which was then solved with the aid of user association and bandwidth allocation algorithms. Multi-stage green energy allocation and multi-BS traffic load balancing algorithms were designed for green energy optimisation in [HA13].

The qualitative comparison of the above-mentioned user association algorithms for hybrid energy powered networks is detailed in Table 2-B.

2.4 Summary and Discussions

This chapter provides an overview of the architecture of HetNets. The existing research outcomes on user association in grid-powered HetNets are surveyed and categorised according to different performance metrics. Since harvesting energy from renewable energy sources could be an attractive solution to improve the overall network energy efficiency, the extensive and detailed review of recent advances on user association in renewable energy powered networks are also presented.

Supporting QoS, spectrum efficiency, energy efficiency and user fairness in 5G networks is an essential requirement for real time applications. How to address these performance metrics at the user association stage is becoming increasingly important [HQ14]. From Table 2-A, it is observed that the most existing works consider part of these performance metrics. As such, more research is required to theoretically analyse the tradeoff

among spectrum efficiency, energy efficiency, QoS and user fairness in HetNets, which can provide more engineering insights regarding the interplay of these performance metrics. In addition, more research work on user association design should be dedicated to address the challenging issues imposed by the inherent nature of HetNets, such as the UL-DL asymmetry, the backhaul bottleneck and the need for efficient mobility support. To make my own contribution to fill the above gap, user association optimisation for grid-powered HetNets is investigated in Chapter 3, which incorporates QoS, spectrum efficiency, energy efficiency and user fairness, as well as addresses the UL-DL asymmetry issue via joint UL and DL user association optimisation.

The deployment of renewable energy sources to supplement conventional power grid for powering BSs indisputably underpins the trend of green communication. However, the intermittent and random nature of renewable energy sources invalidates traditional user association rules designed for conventional cellular networks with constant grid power supply. From Table 2-B, it is observed that the existing work on user association in renewable energy powered networks is quite limited. The principle of user association in such a scenario is far from well understood. In the sequel, more research endeavours should be dedicated in this field, and provide more insights on how to to design the user association policy so as to maximise the utilisation of renewable energy, as well as guarantee the QoS provision in the renewable energy powered networks. To cope with this, user association optimisation for HetNets with solely renewable energy source and hybrid energy sources are investigated in Chapter 4 and Chapter 5, respectively.

Chapter 3

User Association Optimisation for Grid-Powered HetNets

This chapter focuses on conventional grid-powered HetNets, where BSs are able to transmit with constant power and get the continuous energy supply from conventional power grid whenever needed. As summarised in Section 2.4, it is imperative to supporting QoS, spectrum efficiency, energy efficiency and user fairness in the user association design, as well as address the challenging issues imposed by the inherent nature of HetNets, such as the UL-DL asymmetry, the backhaul bottleneck and the need for efficient mobility support. This chapter formulates user association optimisation as a bargaining problem from cooperative game theory, which considers QoS, spectrum efficiency, energy efficiency and user fairness, and also pitches in to resolve the UL-DL asymmetry issue in HetNets. Specifically, in Section 3.1, the general system model and simulation platform used in this chapter are introduced. Then a NBS based user association algorithm is proposed to improve the network DL performance in Section 3.2. Additionally, in order to address the UL-DL asymmetry issue in HetNets, a joint DL and UL user association algorithm using NBS is developed to enhance both DL and UL energy efficiencies in Section 3.3. Finally, taking multi-service into consideration, a NBS based opportunistic user association algorithm is also proposed for QoS provision of the delay constraint

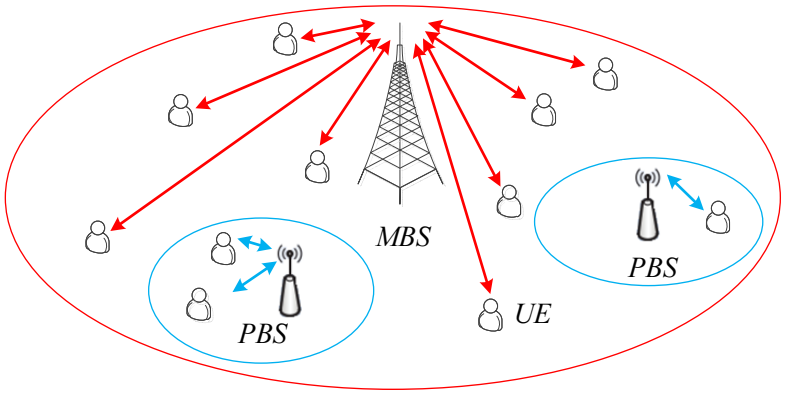


Figure 3.1: Illustration of a 2-tier grid-powered HetNet.

traffic while providing fair resource allocation for the best effort traffic in Section 3.4.

3.1 System Model and Simulation Platform

This section elaborates the system model and simulation platform employed in this chapter.

3.1.1 System Model

The 2-tier HetNets are considered where tier 1 is modelled as macrocell and tier 2 as picocell as shown in Fig. 3.1. There are M BSs, where BS_1 is a MBS, and BS_m is a pico BS (PBS) ($m \in \{2, 3, \cdots, M\}$). All BSs are assumed to share the same frequency band. There are N user equipments (UEs) randomly distributed in HetNets area, denoted as UE_n ($n \in \{1, 2, \cdots, N\}$). Each UE is assumed to be associated with a single BS at any time. Note that the 2-tier model here is able to be extended to the multi-tier model with more than one MBS and more than one type of small BSs.

It is assumed that each user is associated with the same BS for UL and DL transmissions. There are two main techniques to separate UL and DL transmissions on the same physical transmission medium, the time-division duplex (TDD) and frequency-division duplex (FDD). TDD system, with the advantages to accommodate UL and DL traf-

fic asymmetrically, has been regarded as a promising paradigm in 5G networks. In my study, TDD is used to separate DL and UL transmissions with the synchronous operation across the entire network to eliminate BS-to-BS and user-to-user interference [SKEP12]. As an inherent feature of TDD system, channel reciprocity is adopted here, where the UL and DL channels between the same communicating pairs have the same channel power gain [CH12]. g_{mn} is denoted as the average channel power gain between UE_n and BS_m, which includes pathloss and shadowing. $C(\gamma) = \log_2(1+\gamma)$ is used to estimate achievable data rate of UE with a certain user association decision, where γ is the signal-to-interference-plus-noise ratio (SINR) calculated based on the average channel power gain. Note that fast fading is not considered here since the time scale of user association is much larger than the time scale of fast fading [YRC+13].

In order to decouple association and scheduling problem, a Round Robin scheduler is adopted as in [CFM12], where all BSs allocate time-frequency resources among associated UEs equally.

To formulate the user association, the association matrix $\mathbf{X} = [x_{mn}]$ is defined as

$$x_{mn} = \begin{cases} 1, & \text{if } \mathrm{UE}_n \text{ is associated with } \mathrm{BS}_m \\ 0, & \text{otherwise} \end{cases} . \tag{3.1}$$

3.1.1.1 Uplink Model

During the UL open loop control, χ is set as the target signal-to-noise ratio (SNR) at BS. Thus the desired transmit power of UE_n when associated with BS_m is

$$P^u_{mn} = \min\left\{\chi\sigma^2_m / g_{mn}, P^u_{n,\max}\right\}, \tag{3.2}$$

where superscript u here denotes the UL transmission, $P^u_{n,\max}$ is UE_n's maximum transmit power and σ^2_m is the estimated noise level at BS_m.

Then the UL received SINR at BS_m from UE_n is given by

$$\gamma^u_{mn} = \frac{P^u_{mn} g_{mn}}{\sum\limits_{j=1,j\neq n}^{N} \sum\limits_{i=1,i\neq m}^{M} \left(x_{ij} P^u_{ij} / N_i \right) g_{mj} + \sigma^2_m}, \tag{3.3}$$

where σ^2_m is the estimated noise level at BS_m, and $N_i = \sum_{n=1}^{N} x_{in}$ is the number of UEs associated with BS_i. In the UL, the co-channel interference for UE_n comes from the UEs that simultaneously use the same resource block in neighbouring cells. In reality, the UEs that cause interference change from one scheduling cycle to another cycle due to scheduling dynamics. However, in terms of user association, the average interference to a UE is more critical than the instantaneous interference level at each scheduling cycle. Thus in this model, the average interference to a UE is considered. $\sum_{j=1,j\neq n}^{N} \sum_{i=1,i\neq m}^{M} \left(x_{ij} P^u_{ij} / N_i \right) g_{mj}$ is the average interference to UE_n from all UEs in neighbouring cells [CH12].

With the Round Robin scheduler, the UL long-term rate of UE_n when associated with BS_m is

$$r^u_{mn} = W \log_2 \left(1 + \gamma^u_{mn} \right) \Big/ \sum\nolimits_{n=1}^{N} x_{mn}, \tag{3.4}$$

where W is the operating bandwidth.

Then the UL energy efficiency (bits/Joule) of UE_n when associated with BS_m is defined as

$$\eta^u_{mn} = r^u_{mn} / P^u_{mn}. \tag{3.5}$$

For the sake of simplicity, only the transmit power is considered as energy consumption, where the circuit power consumption is omitted, due to the fact that the circuit power consumption is generally independent with the transmit power of BSs or UEs, and normally modelled as a constant value.

3.1.1.2 Downlink Model

Similarly, the DL received SINR at UE_n when associated with BS_m is given by

$$\gamma_{mn}^{d} = \frac{P_m g_{mn}}{\sum_{i=1, i\neq m}^{M} P_i g_{in} + \sigma_n^2}, \tag{3.6}$$

where superscript d here denotes the DL transmission, P_m is the transmit power of BS_m and σ_n^2 is the estimated noise level at UE_n. BS_m is assumed to transmit with P_m all the time, and create interference over the whole bandwidth W.

Then the DL long-term rate of UE_n when associated with BS_m is

$$r_{mn}^{d} = W \log_2\left(1 + \gamma_{mn}^{d}\right) \Big/ \sum_{n=1}^{N} x_{mn}. \tag{3.7}$$

Finally, the DL energy efficiency (bits/Joule) of UE_n when associated with BS_m is

$$\eta_{mn}^{d} = r_{mn}^{d} / P_m. \tag{3.8}$$

3.1.1.3 Fundamentals on NBS

The idea of NBS from cooperative game theory has been widely used for fairly distributing resources among competing players. In [HJL05], NBS was applied for a fair channel allocation in multiuser orthogonal-frequency-division-multiple-access (OFDMA) networks. In [NZ12], NBS was applied into OFDMA cognitive radio (CR) networks, where NBS provided the proportional fairness and efficient power distribution among CR users to maximise the overall throughput of the CR system.

According to cooperative game theory, the bargaining problem is outlined as follows [HJL05]. Assume M players compete for resources and the minimal payoff of each player m $(m \in \{1, 2, \cdots, M\})$ is $U_m^{\min}$, where $\mathbf{U}^{\min} = \left(U_1^{\min}, \cdots, U_m^{\min}, \cdots, U_M^{\min}\right)$.

Assume $\mathbf{U} = (U_1, \cdots, U_m, \cdots, U_M)$ is a closed and convex subset of $\Re^M$ to present the set of feasible payoff allocation that players can get when they cooperate. Since the minimal payoff of each player must be guaranteed, $\left\{U_m \in \mathbf{U} | U_m \geq U_m^{\min}, \forall m \in \{1, 2, \cdots, M\}\right\}$ is a nonempty set. Then $\left(\mathbf{U}, \mathbf{U}^{\min}\right)$ is a M person bargaining problem.

Within the feasible set $\mathbf{U}$, the notion of Pareto optimal is defined as a selection criterion for the bargaining solution.

Definition 3.1: The point $\left(\bar{U}_1, \cdots, \bar{U}_m, \cdots, \bar{U}_M\right)$ is said to be **Pareto optimal**, if and only if there is no other allocation $U_m{}'$ such that $U_m{}' \geq \bar{U}_m, \forall i$, and $U_m{}' > \bar{U}_m, \exists i$, i.e, there exists no other allocation that leads to superior payoff for some players without inferior payoff for some other players.

There might be an infinite number of Pareto optimal points [YMR00]. Further criteria is needed to select a bargaining result. From the perspective of resource sharing, one of the natural criteria is the notion of fairness. This, in general, is a loose term and there are many notions of fairness. One of the commonly used notions is that of max-min fairness which penalises strong players with good conditions, and as a result, generates inferior system performance [HJL05]. In this chapter the NBS fairness, which is inherent from NBS, is adopted. The *proportional fairness* is a special case of NBS fairness [HJL05]. NBS provides a unique and fair Pareto optimal system operation point, which is defined as [HJL05]

Definition 3.2: $\mathbf{U}^*$ is said to be the NBS in $\mathbf{U}$ for $\mathbf{U}^{\min}$, i.e., $\mathbf{U}^* = \phi\left(\mathbf{U}, \mathbf{U}^{\min}\right)$, if the following axioms are satisfied.

1. *Individual Rationality*: $U_m \geq U_m^{\min}, \forall m$.
2. *Feasibility*: $\mathbf{U}^* \in \mathbf{U}$.
3. $\mathbf{U}^*$ is *Pareto optimal*.
4. *Independence of Irrelevant Alternative*: If $\mathbf{U}^* \in \mathbf{U}' \subset \mathbf{U}$, $\mathbf{U}^* = \phi\left(\mathbf{U}, \mathbf{U}^{\min}\right)$, then

$\mathbf{U}^* = \phi\left(\mathbf{U}', \mathbf{U}^{\min}\right)$.

5. *Independence of Linear Transformations*: For any linear scale transformation, φ, $\varphi\left(\phi\left(\mathbf{U}, \mathbf{U}^{\min}\right)\right) = \phi\left(\varphi\left(\mathbf{U}\right), \varphi\left(\mathbf{U}^{\min}\right)\right)$.

6. *Symmetry*: If $\mathbf{U}$ is invariant under all exchanges of players, $\phi_m\left(\mathbf{U}, \mathbf{U}^{\min}\right) = \phi_{m'}\left(\mathbf{U}, \mathbf{U}^{\min}\right), \forall m, m'$.

Axioms 4–6 are the so-called axioms of fairness. Axiom 4 states that the bargaining solution is not affected by enlarging the domain if agrement can be found in a restricted domain. Axiom 5 indicates the bargaining solution is unchanged if the performance objective objectives are affinely scaled. Axiom 6 asserts that the bargaining solution does not depend on the specific label, i.e., players with the same initial points and objective will realise the same performance.

According to [YMR00], there is a unique solution for $\phi\left(\mathbf{U}, \mathbf{U}^{\min}\right)$ that satisfies all six axioms in *Definition 3.2*, and this solution satisfies

$$\begin{aligned} \mathbf{U}* = \arg\max_{\mathbf{U}} & \prod_{m=1}^{M} \left(U_m - U_m^{\min}\right), \\ \text{s.t.} \quad & U_m \geq U_m^{\min}. \end{aligned} \tag{3.9}$$

Based on [YMR00], if U_m is a concave upper-bounded function with convex support, there exists a unique and optimal NBS.

3.1.2 Simulation Platform

To evaluate the performance of proposed algorithms, a HetNet composed of one macrocell and several picocells is simulated.

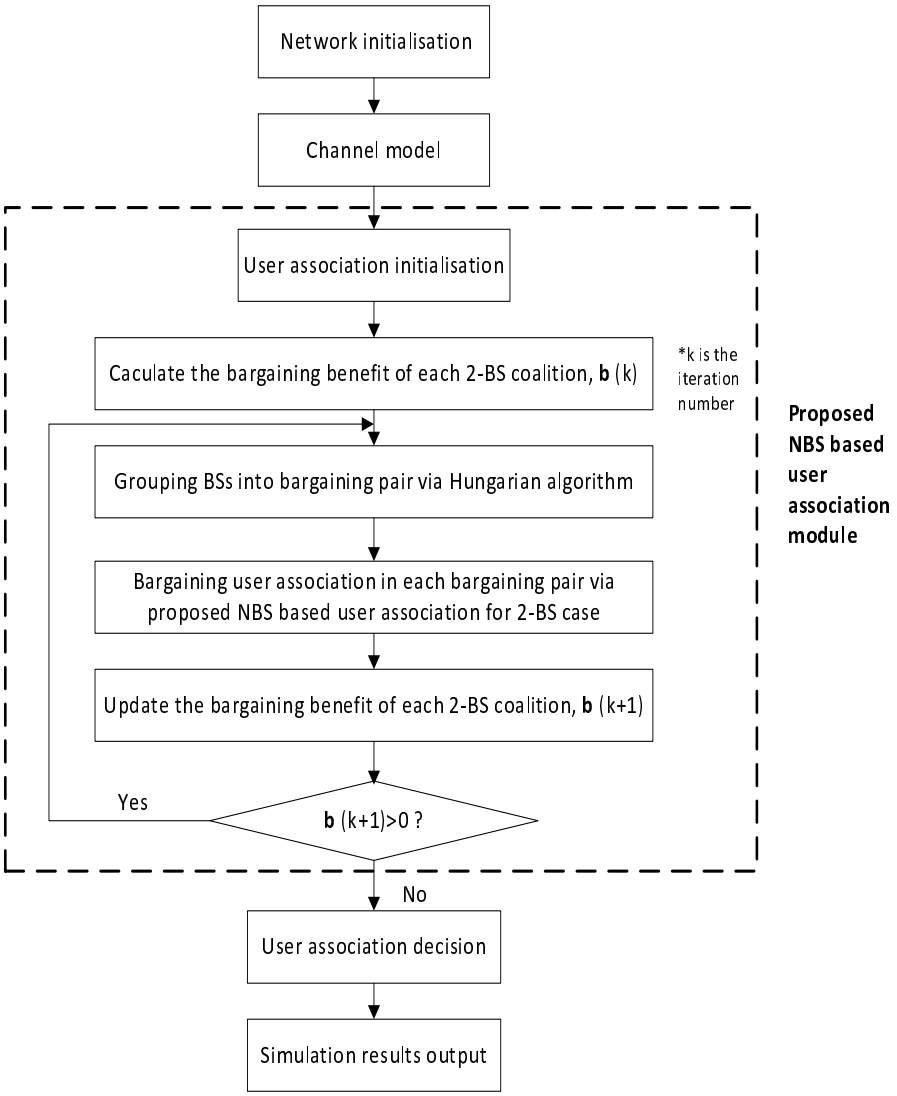


Figure 3.2: Flowchart of simulation platform for user association optimisation in grid-powered HetNets.

3.1.2.1 Simulation Platform Modules

Fig. 3.2 illustrates the flowchart of the simulation platform used in this chapter. The functionality of each module is summarised as follows.

A. Network initialisation

This module initialises the network topology, the user distribution and system parameters.

- Network topology

  To evaluate the performance of proposed algorithms, the 2-tier DL HetNets are simulated. The theoretical analysis throughout this chapter is independent with the spatial distribution and the specific number of BSs. In the simulation, the locations of all BSs are modelled to be fixed. A simple HetNet composed of one macrocell and several picocells is simulated. In the simulation scenario, PBSs are symmetrically located along a circle with radius 120 m and MBS in the centre. The inter-site distance is 500 m. Similar simulation network topology can be found in [Guv11]. It is straightforward to modify the results for different user distributions and different PBS locations.

- User distribution

  In each snap shot of simulation, users are randomly distributed in HetNets geographical area.

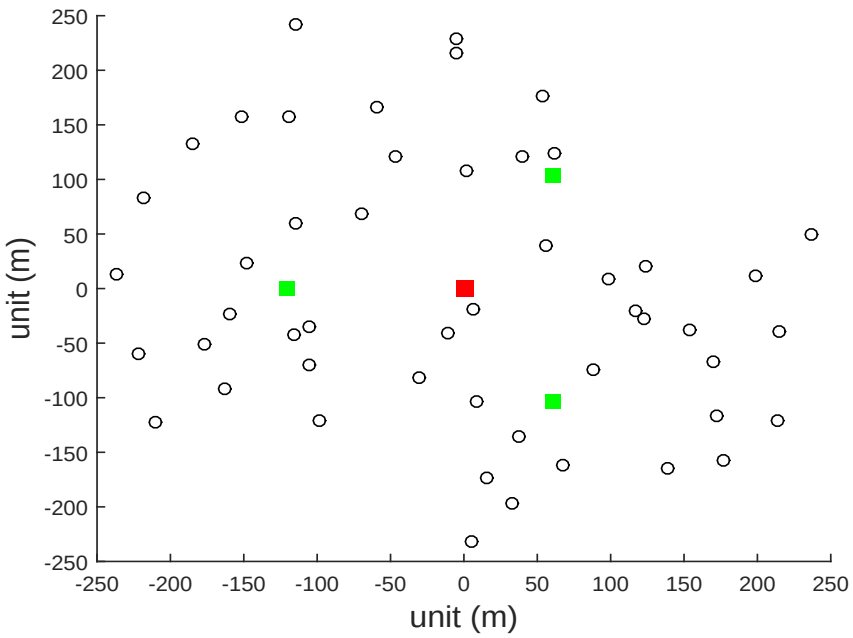


Figure 3.3: Snapshot of the simulated network scenario, where MBS (red square) is located in the centre of the cell, 3 PBSs (green square) are located along the MBS, and 50 users (black circle) are randomly distributed in HetNets geographical area.

Fig. 3.3 shows a snapshot of the simulated network scenario, where MBS (red square) is located in the centre of the cell, 3 PBSs (green square) are located along the MBS, and 50 users (black circle) are randomly distributed in HetNets geographical area.

B. Channel model

This module specifies the channel model in the simulation scenario, and then calculates the channel power gain according to users' and BSs' physical locations.

In HetNets scenario, the channel condition for the transmission between MBS and users is different from that between small BSs and users.

The pathloss models between MBS and user, and between PBS and user in 3GPP standards [3GP10] are adopted in the simulation, which are given by

$$L_{MBS} = 128.1 + 37.6\log_{10} d\,(km)\,, \tag{3.10}$$

$$L_{PBS} = 140.7 + 36.7\log_{10} d\,(km)\,, \tag{3.11}$$

where d is the distance between the BS and the user.

In line with 3GPP standards [3GP10], the log-normal shadow fading with standard deviation (SD) 10dB is also considered in the simulation.

C. Proposed NBS based user association

This module implements proposed NBS based user association algorithms. In Section 3.2.4, the performance of the NBS based user association for DL transmission is evaluated. The performances of NBS based joint UL and DL user association for energy efficient HetNets, and NBS based user association for multi-service HetNets are validated in Section 3.3.4 and Section 3.4.4, respectively.

3.1.2.2 Simulation Parameters

The simulation parameters used in this chapter are summarised as in Table 3-A.

Table 3-A: Simulation Parameters

Parameter	**Value**
Carrier frequency	2 GHz [3GP10]
Inter-site distance	500 m [3GP10]
Bandwidth	10 MHz [3GP10]
Transmit power of MBS	46 dBm
Transmit power of PBS	30 dBm
Maximum transmit power of UE	23 dBm
Noise power density	-174dBm/Hz
Pathloss between MBS and UE	$128.1 + 37.6\log_{10}d\,(km)$ [3GP10]
Pathloss between PBS and UE	$140.7 + 36.7\log_{10}d\,(km)$ [3GP10]
Log-normal shadowing fading SD	10 dB [3GP10]
User required minimal rate	100 Kbps [HJL05]
Number of drops	500

3.2 NBS Based User Association Optimisation in HetNets

This section elaborates the proposed NBS based user association in HetNets, which focuses on the DL transmission only.

3.2.1 Motivation

Game theory, with distinct advantages in investigating the interaction of multiple players, is a powerful tool for solving user association problems. Lines of work [SSV+14, NSMV14, PBS+13] mentioned in Section 2.2.1 formulated the DL user association in HetNets as a many-to-one matching game, which is a non-cooperative game. In non-cooperative game theory modelling, players seek to maximise their own utility, and compete against each other by adopting different strategies. The solution of none-cooperative game is a Nash equilibrium, which may not be Pareto optimal, and may sometimes appear non-rational in a third-person perspective, consequently leading to lower system performance

compared with the cooperative game. Therefore, cooperative game theory modelled user association is worthy of thorough investigation. The bargaining game from cooperative game theory is fairly suitable for modelling the user association problem in HetNets, where BSs are modelled as players to bargain user association for mutual advantages.

In this section a fair user association algorithm is proposed for HetNets, where the user association optimisation is formulated as a bargaining problem. The optimisation objective is to maximise the sum of rate related utility, under users' minimal rate constraints, while considering user fairness and load balancing among cells in different tiers. NBS and coalition are adopted to solve this optimisation problem. Firstly, a two-player bargaining algorithm is developed for two BSs to bargain user association. Then this two-player algorithm is extended to a multi-player bargaining algorithm with the aid of Hungarian algorithm that optimally groups BSs into pairs. Simulation results indicate that the proposed algorithm can effectively offload users from macrocells, improve user fairness, and also achieve comparable sum rate to the algorithm that maximises the sum rate without considering user fairness.

3.2.2 Problem Formulation

The user association problem in HetNets can be modelled as the bargaining problem among different BSs (e.g. MBSs, PBSs) competing for users to connect to them.

The utility of UE_n if associated with BS_m is defined as

$$\mu_{mn} = \log\left(r_{mn}^{d}\right) - \log\left(r_{n}^{\min}\right), \tag{3.12}$$

where $r_n^{\min}$ is the minimal DL data rate required by UE_n.

The logarithm function is widely used to construct utility function [YRC+13]. It is concave and hence has diminishing benefit, which encourages user fairness where allocating more resources to users with lower rate is considered desirable. Furthermore, in this

model, μ_{mn} will be positive only when UE_n's minimal data rate requirement is satisfied.

The payoff of BS_m is defined as the sum utility of all users associated with it, which is given by

$$U_m = \sum_{n=1}^{N} x_{mn} \cdot \mu_{mn}. \tag{3.13}$$

Substituting (3.7) and (3.12) into (3.13), we can get

$$U_m = \sum_{n=1}^{N} x_{mn} \left(\log \left(c_{mn} \Big/ \sum_{n=1}^{N} x_{mn} \right) - \log \left(r_n^{\min} \right) \right), \tag{3.14}$$

where $c_{mn} = W \log_2 \left(1 + \gamma_{mn}^d\right)$.

In the formulated problem, the optimisation goal is to determine which UE n should be associated with BS m, so as to maximise the NBS utility function U, which is given by

$$\begin{aligned} \max_{\mathbf{X}} \quad & U = \prod_{m=1}^{M} \left(U_m - U_m^{\min}\right), \\ s.t. \quad & U_m \geq U_m^{\min}, \ \forall m \\ & x_{mn} = \{0, 1\}, \ \forall m, n \\ & \sum_{m=1}^{M} x_{mn} = 1, \forall n, \end{aligned} \tag{3.15}$$

where $U_m^{\min}$ is the minimal payoff of BS_m, which is set as 0 in order to ensure the data rate of user is no less than the required minimum data rate. $\sum_{m=1}^{M} x_{mn} = 1$ means each user must be associated with a single BS.

The user association problem formulated in (3.15) falls into the class of integer programming problems, where an exact solution usually involves an exhaustive search. However, an exhaustive search is computationally prohibitive when the number of UEs is large. Approximating the problem by its continuous relaxation is a common approach

to the integer programming problem. The idea of continuous relaxation is to enlarge the constraint set to include all convex combinations of the original points. As such, $x_{mn} = \{0, 1\}$ is relaxed to $0 \leq x_{mn} \leq 1$, where x_{mn} specifies the probability that UE_n is associated with BS_m.

Note that, although the probabilistic user association is adopted through continuous relaxation here. The user association proposed in Section 3.2.3 determines the optimal deterministic user association. This will be made clear in the proof of **Proposition 1** in Section 3.2.3.

Therefore, the bargaining problem of user association in HetNets is described as follows. Each BS has U_m as payoff function, and U_m is concave and upper-bounded, since its Hessian matrix is negative semidefinite. The optimisation goal is to determine $\mathbf{X}$ to maximise all U_m simultaneously under the constraint $U_m \geq U_m^{\min}$. It is crucial to design a simple and fast method to find the optimal, unique, and fair $\mathbf{X}$.

3.2.3 NBS Based User Association Algorithm

3.2.3.1 Algorithm for Two-BS Case

This subsection focuses on the two-player bargaining algorithm for two BSs ($M = 2$) which is extended to the multi-player bargaining algorithm for multiple BSs in the next subsection. In this two-player bargaining algorithm, BS_1 and BS_2 are arbitrary parts of BSs in HetNets. They can be modelled as a MBS and a PBS, or two MBSs, or two PBSs. Inspired by the low complexity algorithm in [YC02], the simple two-band partition is applied to determine the user association. The authors of [YC02] focused on subcarrier assignment, and the idea was to sort the order of subcarriers first, and then use a simple two-band partition for subcarrier assignment. In addition, the authors of [YC02] have proven that the two-band partition for two-user subcarrier assignment is near-optimal for the optimisation goal of weighted maximisation of both users' rate.

The proposed two-band UE partition algorithm for NBS based user association can be described in Table 3-B, and in each iteration, ρ_1 and ρ_2 are fixed. Firstly, the initial ρ_1, ρ_2, and NBS utility $U(0)$ are determined based on the initial user association, and we set $i = 1$ (i is the number of iteration). Then all UEs are sorted in decreasing order according to $\left(r_{1n}^d\right)^{\rho_1} / \left(r_{2n}^d\right)^{\rho_2}$. The first UE$_1$,..., UE$_n$, $n \in \{1, 2, \cdots, N-1\}$, are arranged to connect to the BS$_1$, and the rest of users connect to the BS$_2$. Every $U(n)$ is calculated, and the partition $\aleph = \arg\max_n U(n)$ is selected. $U_{\max}(i) = U(\aleph)$ and $i = i + 1$ are set. The UE partition process is repeated with the updated ρ_1 and ρ_2 according to the new partition $\aleph$, until no more improvement can be achieved for the NBS utility U.

Table 3-B: Two-band UE Partition for NBS Based User Association

Step 1. Initialisation Initialise the user association, such that BS$_m$'s minimal payoff can be guaranteed. Calculate $U_{\max}(0) = U(0)$, ρ_1 and ρ_2. Set $i = 1$.
Step 2. Sort UEs Arrange the UE index from the largest to smallest according to $\left(r_{1n}^d\right)^{\rho_1} / \left(r_{2n}^d\right)^{\rho_2}$.
Step 3. for $n = 1, \cdots, N-1$ UE$_1$,..., UE$_n$ connect to the BS$_1$; UE$_{n+1}$,..., UE$_N$ connect to the BS$_2$; Calculate $U(n)$. **end for**
Choose the two-band partition $\aleph = \arg\max_n U(n)$ **which generates the largest** U **satisfying the constraints.** Set $U_{\max}(i) = U(\aleph)$.
Step 5. Update user association **if** $U_{\max}$ cannot be increased by updating ρ_1 and ρ_2, the iteration ends; **otherwise**, update ρ_1 and ρ_2 according to the new partition, set $i = i + 1$, and then go to ***step 2***.

This two-band UE partition algorithm has the complexity of $O\left(N^2\right)$ for each iteration. The binary search algorithm can further improve it with complexity $O(N\log_2 N)$ for each iteration. According to the simulation, this two-band UE partition algorithm converges within three rounds.

Proposition 1: The two-band UE partition shown in Table 3-B is near-optimal to the optimisation problem in (3.15) with $M = 2$.

Proof. In the NBS based user association, the two-player game is to maximise the NBS

utility $U = \left(U_1 - U_1^{\min}\right)\left(U_2 - U_2^{\min}\right)$. As the constraint $x_{mn} = \{0, 1\}$ is relaxed to continuous value with $0 \leq x_{mn} \leq 1$. Based on (3.15), the Lagrangian function as a function of x_{mn} can be written as

$$L = \prod_{m=1}^{2}\left(U_m - U_m^{\min}\right) + \sum_{n=1}^{N}\lambda_n\left(\sum_{m=1}^{2} x_{mn} - 1\right), \tag{3.16}$$

where λ_n is the Lagrangian multiplier. By taking the Karush-Kuhn-Tucker (KKT) condition, and substituting (3.14) into (3.16), the derivative of (3.16) with respect to x_{mn} is

$$\frac{\log\left(c_{1n} \Big/ \sum_{n=1}^{N} x_{1n}\right) - 1 - \log\left(r_n^{\min}\right)}{U_1 - U_1^{\min}} = \frac{\log\left(c_{2n} \Big/ \sum_{n=1}^{N} x_{2n}\right) - 1 - \log\left(r_n^{\min}\right)}{U_2 - U_2^{\min}}. \tag{3.17}$$

Since $r_{mn}^d = c_{mn} \Big/ \sum_{n=1}^{N} x_{mn}$, the equation (3.17) can be rewritten as

$$\frac{\log\left(r_{1n}^d\right) - 1 - \log\left(r_n^{\min}\right)}{U_1 - U_1^{\min}} = \frac{\log\left(r_{2n}^d\right) - 1 - \log\left(r_n^{\min}\right)}{U_2 - U_2^{\min}}. \tag{3.18}$$

The left side of equation (3.18) can be interpreted as the marginal benefit for BS_1 if UE_n connects to it. The right side is the marginal benefit for BS_2 if UE_n connects to it. If UE_n is associated with both BS_1 and BS_2, (3.18) needs to be satisfied with equality. When UE_n only connects to one BS, the equation (3.18) becomes inequality. If the left side of equation (3.18) is greater than the right side, UE_n should connect to BS_1 and vice versa to BS_2. Function f is defined as the difference of left and right sides of equation (3.18).

$$f\left(\frac{\left(r_{1n}^d\right)^{\rho_1}}{\left(r_{2n}^d\right)^{\rho_2}}\right) = \log\left(\frac{\left(r_{1n}^d\right)^{\rho_1}}{\left(r_{2n}^d\right)^{\rho_2}}\right) - \log\left(\frac{\left(r_n^{\min}\right)^{\rho_1}}{\left(r_n^{\min}\right)^{\rho_2}}\right) - \rho_1 + \rho_2, \tag{3.19}$$

where

$$\rho_1 = \left(U_1 - U_1^{\min}\right)^{-1}, \tag{3.20}$$

and

$$\rho_2 = \left(U_2 - U_2^{\min}\right)^{-1}. \tag{3.21}$$

Here $f\left(\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho_2}\right)$ can be interpreted as the marginal benefit difference of BS_1 and BS_2 when UE_n is associated with them. When $f\left(\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho_2}\right) > 0$, UE_n should be associated with BS_1, and when $f\left(\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho 2}\right) < 0$, UE_n should be associated with BS_2. It is obvious that $f\left(\left(r_{1n}^d\right)^{\rho 1}\Big/\left(r_{2n}^d\right)^{\rho_2}\right)$ is a monotonic function with $\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho_2}$. Then the indexes of UEs are arranged to make $\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho_2}$ decrease in n. With the fixed ρ_1 and ρ_2, $f\left(\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho_2}\right)$ is a monotonic function of n, which is similar to the weighted maximisation function in [YC02], so the two-band partition is the near-optimal solution.

Within each iteration, ρ_m is fixed, and the two-band UE partition achieves the near-optimal solution. Then in the next iteration, ρ_m is updated. Remember U_m^* is the NBS. If $U_1 > U_1^*$ and $U_2 < U_2^*$, ρ_1 is small and ρ_2 is relatively large. Consequently, the marginal benefit of BS_1 will be reduced, leading to a disadvantage for bargaining user association in the next iteration, and vice versa. This is one explanation why the proposed two-band UE partition converges to the NBS [HJL05].

It is worthy mentioning that since $x_{mn} = \{0, 1\}$ is relaxed to $0 \leq x_{mn} \leq 1$, in the end, UE_n with $f\left(\left(r_{1n}^d\right)^{\rho_1}/\left(r_{2n}^d\right)^{\rho_2}\right) = 0$ will be associated with BS_1 and BS_2 simultaneously. This UE_n can be named as boundary UE, and all the other UEs will connect to either BS_1 or BS_2. As similarly stated in [YC02], given the number of UEs is much larger than the number of BSs, this boundary UE can be associated with either BS arbitrarily without affecting the system performance. □

3.2.3.2 Algorithm for Multi-BS Case with Coalition

For the case where there are more than two BSs, the most work in literature focused on solving the user association problem among multiple BSs together in a centralised method [CFM12]. Since the problem itself is combinatorial, this centralised method such as brute force will incur high computational complexity with $O\left(M^N\right)$ (M and N are number of BSs and UEs, respectively), such computation is essentially impossible for even medium-sized HetNets. In this section, a two-step iterative algorithm is proposed. Firstly, BSs are grouped into pairs called coalitions. Then in each coalition, the two-player bargaining algorithm in Table 3-B is executed. BSs are regrouped and re-bargaining until convergence. In this way, the computational complexity can be greatly reduced.

Grouping BSs into pairs is an assignment problem. The most common way to solve this problem is the random method, where BSs are randomly grouped into pairs to bargain the user association. If the convergence speed is measured by the round of bargaining, this random method may converge slowly, since the bargaining in the randomly grouped pair will have little utility improvement over the utility before bargaining. In order to reduce the computational complexity and speed up the convergence, the Hungarian algorithm [PS98] can be used to solve such a assignment problem. This assignment problem will be modelled in detail as below.

The benefit for ith BS bargaining with jth BS is defined as b_{ij}, which is the element of the matrix $\mathbf{b}$,

$$b_{ij} = \max\left(U\left(\tilde{U}_i, \tilde{U}_j\right) - U\left(\widehat{U}_i, \widehat{U}_j\right), 0\right), \tag{3.22}$$

where $\tilde{U}_i$ and $\tilde{U}_j$ are the ith BS's and jth BS's payoff after bargaining, $\widehat{U}_i$ and $\widehat{U}_j$ are the ith BS's and jth BS's payoff before bargaining. Obviously, $b_{ii} = 0\,, \forall i$ and $\mathbf{b}$ is symmetric. The proposed algorithm in Table 3-B can calculate each $b_{ij}\,, \forall i, j$. The computational complexity is about $O\left(M^2 N \log_2 N\right)$.

The coalition assignment matrix $\mathbf{h} = [h_{ij}]$ is defined as

$$h_{ij} = \begin{cases} 1, & \text{if } \mathrm{BS_i} \text{ bargains with } \mathrm{BS_j} \\ 0, & \text{otherwise} \end{cases}. \tag{3.23}$$

Obviously, $\mathbf{h}$ is symmetric.

The assignment problem is how to group the bargaining pair, in order to maximise the overall benefit, which is given by

$$\begin{aligned} &\max_{\mathbf{h}} \sum_{i=1}^{M} \sum_{j=1}^{M} h_{ij} b_{ij}, \\ &s.t. \quad \sum_{i=1}^{M} h_{ij} = 1, \ \forall j \\ &\qquad \sum_{j=1}^{M} h_{ij} = 1, \ \forall i \\ &\qquad h_{ij} \in \{0, 1\}, \quad \forall i, j. \end{aligned} \tag{3.24}$$

Due to the minimisation goal of the Hungarian algorithm, the optimisation goal of (3.24) can be modified as $\min_{\mathbf{h}} \sum_{i=1}^{M} \sum_{j=1}^{M} h_{ij} B_{ij}$, where $B_{ij} = -b_{ij}$. The procedure of the Hungarian algorithm can be described as Table 3-C.

Table 3-C: The Procedure of Hungarian Algorithm

1. Subtract the row minimum from the entries in each row of **B** –each row has at least one zero.
2. Subtract the column minimum from the entries in each column of **B** –each row and each column have at least one zero.
3. Select rows and columns, and draw line across them, –all zeros are covered and the number of lines is minimum.
4. If the number of lines = M select a combination and the sum of such combination is zero. If the number of lines < M go to ***step 5***.
5. Find the smallest entry that is not covered by any line, –subtract it from each entry which is not covered by lines, –add it to the each entry which is covered by lines, –go to ***step 3***.

The NBS based user association algorithm for multi-BS is shown in Table 3-D, where in each iteration, the NBS utility $U = \prod_{m=1}^{M} \left(U_m - U_m^{\min}\right)$ is nondecreasing in step 2 and step 3, and the NBS utility is upper bounded. Consequently, the proposed algorithm is

convergent.

Table 3-D: NBS Based User Association Algorithm for Multi-BS

Step 1. Initialise the user association
Associate all users to BSs.
Step 2. Group coalition
If the number of BSs is odd, a dummy BS is created, and no
BS bargains user association with this dummy BS.
-*Random Method*: Group the 2-BS coalition randomly.
-*Hungarian Algorithm*: The coalition is grouped by the algorithm in Table 3-C
Step 3. Bargain in each coalition
Bargain the user association by the algorithm in Table 3-B
Step 4. Repeat
Go back to ***Step 2***, until no further improvement can be achieved,
i.e., **b** is equal to zero matrix.

The complexity of Hungarian algorithm is $O\left(M^3\right)$, therefore the overall complexity for each round of the proposed NBS based algorithm for multi-BS is $O\left(M^2 N \log_2 N + M^3\right)$. Moreover, the simulation result in the next subsection shows that the proposed algorithm mostly converges within six rounds.

3.2.4 Simulation Results and Conclusions

3.2.4.1 Simulation Results

The performance of the proposed NBS based user association algorithm (NBS algorithm) is compared with the conventional max RSS algorithm, CRE algorithm [Guv11], and maximum sum rate based association algorithm (max sum rate algorithm) [CFM12]. In the max RSS algorithm, users are associated with the BS from which they receive the highest DL RSS. In the CRE algorithm, a positive bias is added to the DL RSS from PBS, and users are associated with the BS from which they receive the highest biased RSS. The bias of CRE algorithm is set as 10 dB in the simulation. In the max sum rate algorithm in [CFM12], users are associated with BSs in order to maximise the sum rate of all users.

Assuming 40 users are randomly distributed within HetNets area, Fig. 3.4 shows the average ratio of users associated with PBSs with different numbers of PBSs in HetNets

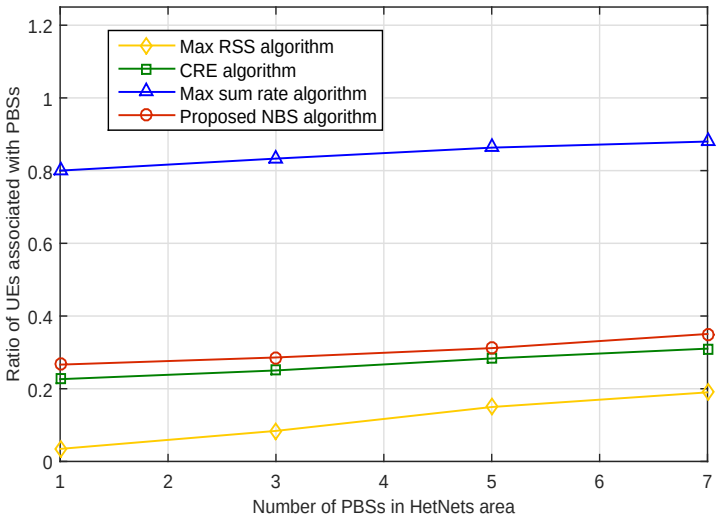


Figure 3.4: Ratio of UEs associated with PBSs versus different numbers of PBSs in HetNets area.

area. It is obvious that with the increasing number of available PBSs, more users will be associated with PBSs. Due to the bias added to the received RSS from the PBS, the CRE algorithm accommodates more users in picocells compared to the max RSS algorithm. The proposed NBS algorithm even offloads more users from the macrocell to picocells compared with the previous two algorithms, since the current cell load is taken into consideration when deciding which BS a certain user will be associated with. The max sum rate algorithm allocates the majority of users to PBSs. This is because in order to maximise sum rate of all users, only the users closer to the MBS will occupy the macrocell, and the rest users will be forced to connect to picocells. Consequently, the load is very imbalanced between MBSs and PBSs, and users associated with PBSs will not get the required data rate. In the following simulation, the number of PBSs in HetNets area is set as 3.

Fig. 3.5 shows the user fairness in these four algorithms with the increase of user number in HetNets area. The widely accepted Jain's fairness index (JFI) [BJ07] is used

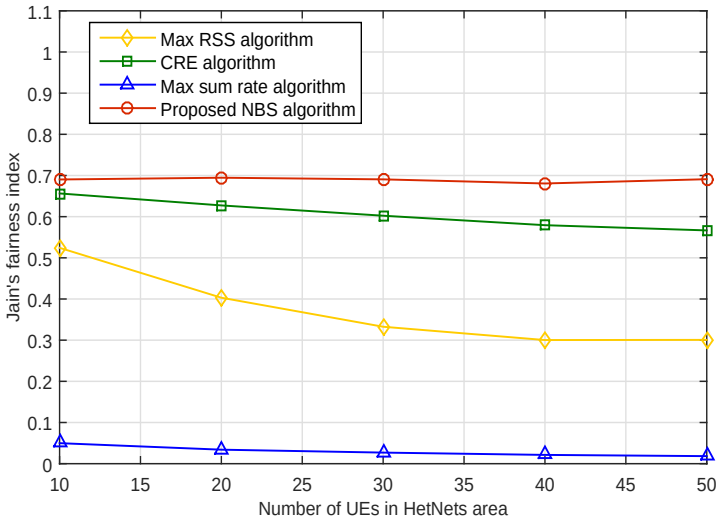


Figure 3.5: Jain's fairness index versus different numbers of UEs in HetNets area.

to evaluate the user fairness, which is defined as

$$J = \frac{\left(\sum_{n=1}^{N} r_n\right)^2}{N \sum_{n=1}^{N} r_n^2}, \tag{3.25}$$

where r_n is data rate of UE_n when the user association is determined.

The larger J indicates better fairness among users. From the Fig. 3.5, it is obvious that the proposed algorithm is superior to the other algorithms in terms of the fairness. This indicates that in the proposed NBS algorithm, users can fairly access resources. Furthermore, Fig. 3.5 illustrates that the number of users does not affect the user fairness in the proposed algorithm. Nevertheless, the JFI values of max RSS and CRE algorithm decrease with the user growth, which means the more users in HetNets area, the more obvious the discrimination of data rate among users is. The JFI value of max sum rate algorithm is lowest compared with the other three algorithms. This can be explained by the fact that in the max sum rate algorithm, the users closer to the MBS will achieve large transmission rate, while the rest of users are forced to connect to picocells and have low data rate.

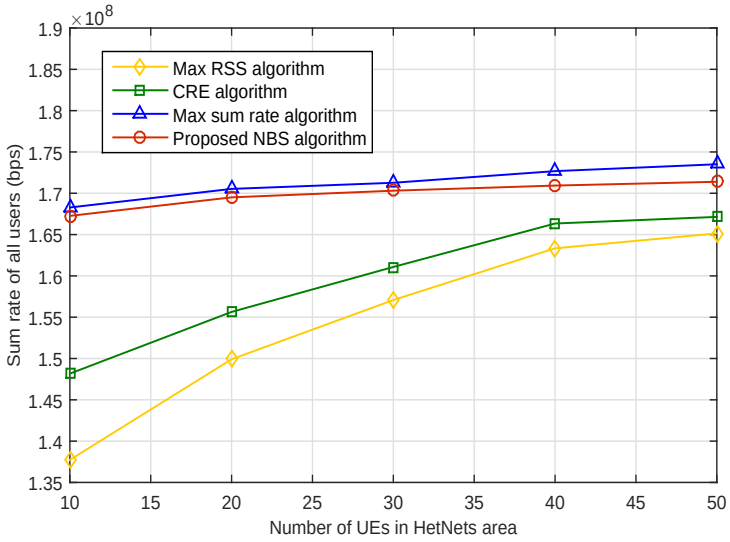


Figure 3.6: Sum rate of all users versus different numbers of UEs in HetNets area.

Fig. 3.6 is the sum rate of all users in these four algorithms with the increase of user number in HetNets area. It is observed that all four algorithms have better performances with the increasing number of users, especially the max RSS and CRE algorithms. This is due to the fact that the increasing number of users will lead to higher probability that some users will be in the vicinity of PBS, which may result in more users being associated with PBSs, thereby benefiting the load balancing among BSs, and enhancing the sum rate of all users. In addition, the performance improvement gets saturated gradually. The proposed NBS algorithm has a similar performance with the max sum rate algorithm, and has a better performance than max RSS and CRE algorithms. Although the max sum rate algorithm slightly outperforms the proposed NBS algorithm, the high sum rate of all users in the max sum rate algorithm is achieved at the cost of low user fairness.

Fig. 3.7 shows the cumulative distribution function (CDF) of rounds number that is necessary for the proposed NBS algorithm convergence by using random method and Hungarian algorithm, respectively. The proposed algorithm with Hungarian algorithm converges in about one to six rounds, while the proposed algorithm with random method converges much slower. This is because Hungarian algorithm can find the best bargaining

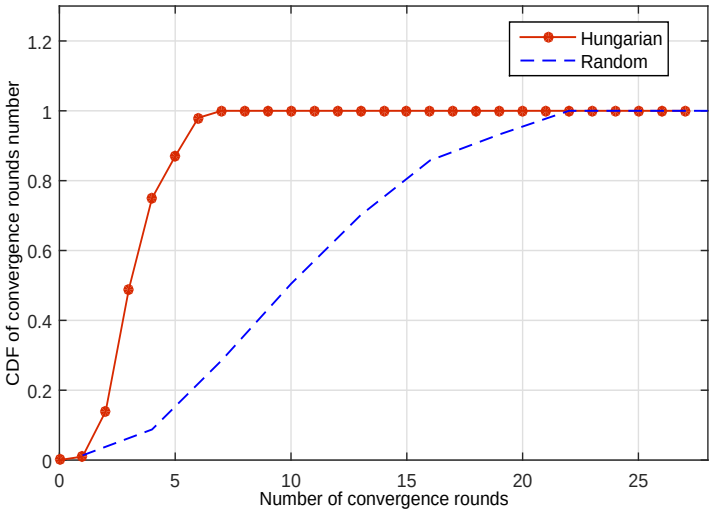


Figure 3.7: CDFs of convergence rounds.

pairs, resulting in higher convergence rate, lower computational complexity and less computational cost.

3.2.4.2 Conclusions

In this section, NBS from cooperative game theory, was applied to develop a fair user association algorithm for HetNets. In the proposed algorithm, user association optimisation problem was formulated as a bargaining problem, which aims to maximise the sum of rate related utility. Load balancing among cells in different tiers, user fairness and users' minimal data rate requirements were taken into consideration. Firstly, the bargaining algorithm for two BSs was presented, and then this algorithm was extended to a multi-player bargaining algorithm for multiple BSs with the aid of Hungarian algorithm. The proposed algorithm has a low computational complexity of $O\left(M^2 N \log_2 N + M^3\right)$ for each iteration. Simulation results indicate that the proposed algorithm can effectively offload users from macrocell to picocells, improve user fairness and achieve comparable sum rate of all users to the max sum rate algorithm.

3.3 Joint UL and DL User Association for Energy-Efficient HetNets Using NBS

This section elaborates the proposed Joint Uplink and Downlink User Association (JUDUA) that takes both UL and DL energy efficiencies into consideration when deciding the serving BS for UEs.

3.3.1 Motivation

Most of the research on user association in HetNets investigated the problem from either the DL or the UL perspective. However, HetNets introduce a major asymmetry between UL and DL in terms of transmit power, traffic and hardware limitations. Due to the UL-DL asymmetry in HetNets, the user association that is optimal for the DL or UL only will not be effective for the opposite direction. Specifically, the conventional max DL RSS based user association rule may associate the user with the far-away marcocell rather than the nearby small cell. As a result, this user will generate high UL interference to the users in nearby picocells. As illustrated in Fig. 3.8, the UE is located in the vicinity of a PBS. Despite the distance from the UE to the PBS is shorter than that from the UE to the MBS, this UE is associated with the MBS due to the stronger DL RSS from the MBS. In the UL, this UE has to transmit at a potentially excessive power for guaranteeing the target received signal strength in the UL, thereby inflicting a high UL interference on the small cell users, hence degrading both the spectrum and energy efficiencies as well as shortening the battery recharge period.

On the other hand, given energy efficiency is one of the important requirements for 5G networks, there have been some research works on energy-aware user association in HetNets. However, most of them only focused on the DL performance and did not take the UL performance into account. To the best of our knowledge, only a limited number of existing works have considered the joint UL and DL user association in HetNets.

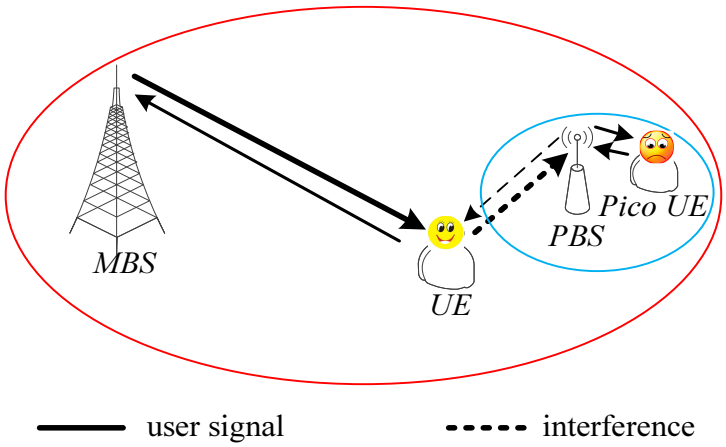


Figure 3.8: DL RSS based user association causes severe UL interference.

In [CH12], an user association algorithm was proposed, and the objective function was formulated as the weighted difference between number of accepted UEs and sum of UL transmit power. However, the algorithm performance highly depended on the chosen weight, which was heuristically obtained in [CH12]. In this section, joint UL and DL User Association (JUDUA) is proposed to maximise the sum of log-scale UL and DL energy efficiencies among all UEs.

3.3.2 Problem Formulation

In JUDUA, BSs are modelled as players in the bargaining problem, and the payoff of BS_m is defined as the sum utility of all the UEs associated with it

$$U_m = \sum_{n=1}^{N} x_{mn}\mu_{mn}, \tag{3.26}$$

where u_{mn} is the utility of UE_n when associated with BS_m.

For JUDUA, the UL energy efficiency of UE should improve along with the DL energy efficiency, so the utility of UE should combine both UL and DL energy efficiencies. Furthermore, in order to preserve some degree of fairness and avoid any BS or UE being starved, the energy efficiency proportional fairness criterion as in [GSSBH11] is applied

in the problem formulation. As such, the utility of UE_n when associated with BS_m is defined as

$$\mu_{mn} = \log\left(\eta^u_{mn}\right) + \log\left(\eta^d_{mn}\right). \tag{3.27}$$

Then the optimisation problem of JUDUA is formulated as

$$\begin{aligned} \max_{\mathbf{X}} \quad & U = \prod_{m=1}^{M} \left(U_m - U_m^{\min}\right), \\ \text{s.t.} \quad & U_m \geq U_m^{\min}, \ \forall m \\ & x_{mn} = \{0, 1\}, \ \forall m, n \\ & \textstyle\sum_{m=1}^{M} x_{mn} = 1, \ \forall n, \end{aligned} \tag{3.28}$$

where U is the NBS utility, and $U_m^{\min}$ is the minimal payoff of BS_m, which is set as 0 indicating the case when no UE is associated with BS_m. $\sum_{m=1}^{M} x_{mn} = 1$ indicates each UE must be associated with a single BS.

The user association problem formulated in (3.28) falls into the class of integer programming problems, where an exact solution usually involves an exhaustive search. However, an exhaustive search is computationally prohibitive when the number of UEs is large. Approximating the problem by its continuous relaxation is a common approach to the integer programming problem. The idea of continuous relaxation is to enlarge the constraint set to include all convex combinations of the original points. As such, $x_{mn} = \{0, 1\}$ is relaxed to $0 \leq x_{mn} \leq 1$, where x_{mn} specifies the probability that UE_n is associated with BS_m.

Note that, although the probabilistic user association is adopted through continuous relaxation here. The user association proposed in Section 3.3.3 determines the optimal deterministic user association. This will be made clear in the proof of **Proposition 2** in Section 3.3.3.

To sum up, the bargaining problem of JUDUA in energy-efficient HetNets is described

as follows. Each BS_m has the payoff U_m, which is concave and upper-bounded, since its Hessian matrix is negative semidefinite. The optimisation goal is to determine $\mathbf{X}$ to maximise all U_m simultaneously under the constraint $U_m \geq U_m^{\min}$. It is crucial to design a simple and fast method to find the optimal, unique and fair $\mathbf{X}$.

3.3.3 Joint UL and DL User Association Algorithm

3.3.3.1 Algorithm for Two-BS Case

The same as the Section 3.2.3.1, this subsection focuses on the two-player bargaining algorithm for two BSs ($M = 2$). Then the next subsection proceeds with the multi-player bargaining algorithm for multiple BSs. In the two-player bargaining algorithm, BS_1 and BS_2 are arbitrary parts of BSs in HetNets. They can be modelled as a MBS and a PBS, or two MBSs, or two PBSs. Inspired by the low complexity algorithm in [YC02], the so-called two-band partition is applied to JUDUA. It was shown in [YC02] that the two-band partition is near-optimal for the optimisation goal of weighted rate maximisation.

The proposed two-band UE partition for JUDUA is described in Table 3-E, and in each iteration, ρ_1 and ρ_2 are fixed. Firstly, the initial ρ_1, ρ_2, and NBS utility $U(0)$ are determined based on the initial user association, and we set $i = 1$ (i is the number of iteration). Then all UEs are sorted in decreasing order according to $(\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}$. The first UE_1,..., UE_n, $n \in \{1, 2, \cdots, N-1\}$, are arranged to be associated with BS_1, and the rest of UEs are associated with BS_2. Every $U(n)$ is calculated, and the partition $\aleph = \arg\max_n U(n)$ is selected. $U_{\max}(i) = U(\aleph)$ and $i = i + 1$ are set. The UE partition process is repeated with the updated ρ_1 and ρ_2 according to the new partition $\aleph$, until no more improvement can be achieved for NBS utility U.

This two-band UE partition has the complexity of $O\left(N^2\right)$ for each iteration, which can be further improved by the binary search algorithm with a complexity of $O(N\log_2 N)$. According to simulations, this two-band UE partition converges within three rounds.

Table 3-E: Two-band UE Partition for JUDUA

Step 1. Initialisation
Initialise user association and guarantee BS_m's minimal payoff.
Set $U_{\max}(0) = U(0)$, and calculate ρ_1 and ρ_2.
Set $i = 1$.
Step 2. Sort UEs
Sort UEs from largest to smallest according to $(\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}$.
Step 3. for $n = 1, \cdots, N-1$
Calculate $U(n)$, where UE_1,..., UE_n are associated with BS_1,
and UE_{n+1},..., UE_N are associated with BS_2.
end for
Step 4. Choose the two-band partition $\aleph = \arg\max_n U(n)$
which generates the largest U **satisfying the constraints.**
Set $U_{\max}(i) = U(\aleph)$.
Step 5. Update user association
if $U_{\max}$ cannot be increased by updating ρ_1 and ρ_2, the iteration ends;
otherwise, update ρ_1 and ρ_2 according to the new partition,
set $i = i + 1$, and then go to ***step 2***.

Proposition 2: The two-band UE partition shown in Table 3-E is near-optimal to the optimisation problem in (3.28) with $M = 2$.

Proof. In JUDUA, the two-player bargaining is to maximise the NBS utility $U = \left(U_1 - U_1^{\min}\right)\left(U_2 - U_2^{\min}\right)$. As the constraint $x_{mn} = \{0, 1\}$ is relaxed to continuous values with $0 \leq x_{mn} \leq 1$, the Lagrangian function of (3.28) as a function of x_{mn} is

$$L = \prod_{m=1}^{2}\left(U_m - U_m^{\min}\right) + \sum_{n=1}^{N}\lambda_n\left(\sum_{m=1}^{2}x_{mn} - 1\right), \tag{3.29}$$

where λ_n is the Lagrangian multipliers. By taking the KKT condition and substituting (3.26), (3.27) into (3.29), the derivative of (3.29) with respect to x_{mn} is

$$\frac{\log\left(\eta_{1n}^u\right) + \log\left(\eta_{1n}^d\right) - 2}{U_1 - U_1^{\min}} = \frac{\log\left(\eta_{2n}^u\right) + \log\left(\eta_{2n}^d\right) - 2}{U_2 - U_2^{\min}}. \tag{3.30}$$

The left and right side of (3.30) can be interpreted as the marginal benefits of UE_n for BS_1 and BS_2, respectively. When UE_n is only associated with one BS, equation (3.30) becomes inequality. If the left side of (3.30) is greater than the right side, UE_n should be associated with BS_1 and vice versa with BS_2. Taking the difference between left and right sides of (3.30), function f is defined as the marginal benefit difference between BS_1

and BS_2 when UE_n is associated with them

$$f\left(\frac{(\eta_{1n})^{\rho_1}}{(\eta_{2n})^{\rho_2}}\right) = \log\left(\frac{(\eta_{1n})^{\rho_1}}{(\eta_{2n})^{\rho_2}}\right) - 2\rho_1 + 2\rho_2, \tag{3.31}$$

where $\eta_{1n} = \eta_{1n}^u \eta_{1n}^d$, $\eta_{2n} = \eta_{2n}^u \eta_{2n}^d$, $\rho_1 = \frac{1}{U_1 - U_1^{\min}}$, and $\rho_2 = \frac{1}{U_2 - U_2^{\min}}$.

Whether UE_n should be associated with BS_1 or BS_2 can be determined by checking whether $f\left((\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}\right)$ is greater or less than zero. It is obvious that $f\left((\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}\right)$ is a monotonic function with $(\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}$. Then the indexes of UEs are sorted to make $(\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}$ decrease in n. With fixed ρ_1 and ρ_2, $f\left((\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}\right)$ is a monotonic function of n. Then function (3.31) is similar to the weighted maximisation function in [YC02], so the two-band partition is the near-optimal solution.

Within each iteration, ρ_m is fixed, and the two-band UE partition achieves the near-optimal solution. Then in the next iteration, ρ_m is updated. Remember U_m^* is the NBS. If $U_1 > U_1^*$ and $U_2 < U_2^*$, ρ_1 is small and ρ_2 is relatively large. Consequently, the marginal benefit of BS_1 will be reduced, leading to a disadvantage for bargaining user association in the next iteration, and vice versa. This is one explanation why the proposed two-band UE partition for JUDUA converges to the NBS [HJL05].

It is worth mentioning that due to the continuous relaxation of x_{mn}, in the end, the UE_n with $f\left((\eta_{1n})^{\rho_1}/(\eta_{2n})^{\rho_2}\right) = 0$ will be associated with BS_1 and BS_2 simultaneously, this UE_n can be named as boundary UE, and all the other UEs are associated with either BS_1 or BS_2. Similar to the explanation in [YC02], given the number of UEs is much larger than the number of BSs, the boundary UE can be associated with either BS arbitrarily without affecting the system performance. □

3.3.3.2 Algorithm for Multi-BS Case with Coalition

The procedure of the algorithm for multi-BS case with coalition is the same as that in Section 3.2.3.2.

3.3.4 Simulation Results and Conclusions

3.3.4.1 Simulation Results

In the simulation, three PBSs are deployed around one MBS. The performance of JUDUA is compared with the conventional max RSS algorithm and CRE algorithm [Guv11]. In the max RSS algorithm, UEs are associated with the BS from which they receive the highest DL RSS. In the CRE algorithm, a positive bias is added in the DL RSS from PBS, and UEs are associated with the BS from which they receive the highest biased DL RSS. In the simulation, the bias is set as 10 dB, and the target received SNR at BS is set as 10 dB as well [CH12].

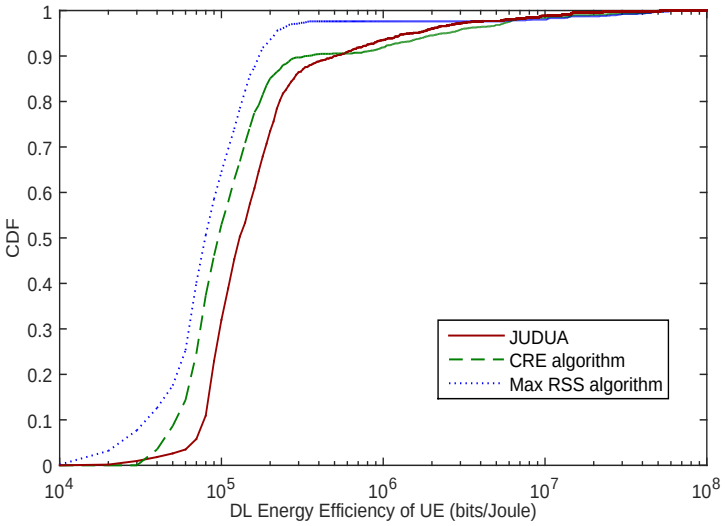


Figure 3.9: CDFs of UE DL energy efficiency.

Assuming 30 UEs are randomly distributed in HetNets area, Fig 3.9 and Fig 3.10 are CDFs of UE DL and UL energy efficiencies in HetNets, respectively. Due to the joint optimisation of UL and DL energy efficiencies in JUDUA, both for the UL and DL, the CDFs of JUDUA improve significantly at low energy efficiency versus CDFs of max RSS algorithm and CRE algorithm. The CDF of CRE algorithm catches up at the DL energy efficiency of $8 * 10^5$ bits/Joule, and at the UL energy efficiency of $3 * 10^9$

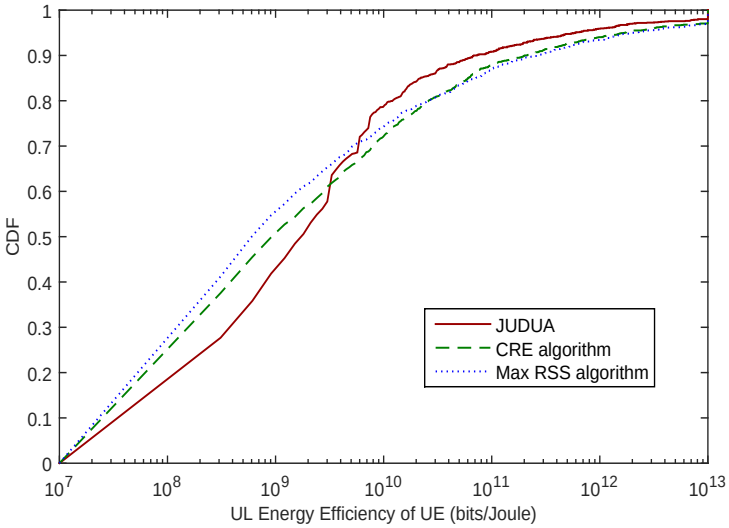


Figure 3.10: CDFs of UE UL energy efficiency.

bits/Joule, respectively. This is because in JUDUA, the energy efficiency proportional fairness criterion provides a more uniform energy efficiency by taking resources from strong UEs. As such, JUDUA improves the user fairness in terms of energy efficiency compared with the other algorithms.

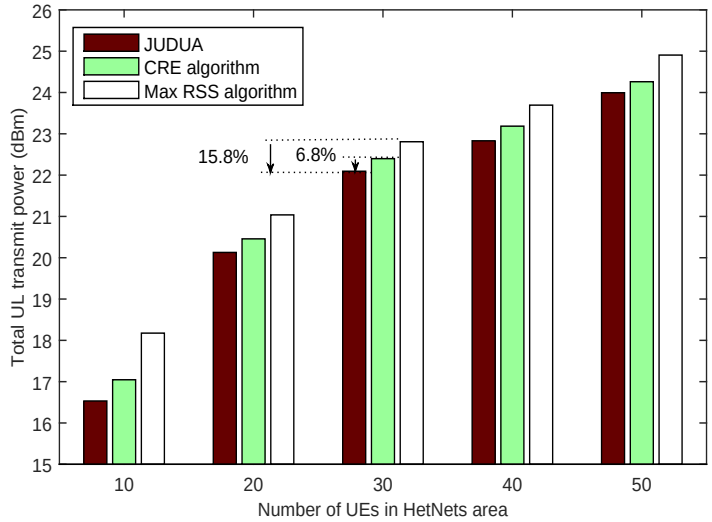


Figure 3.11: Total UL transmit power versus different numbers of UEs in HetNets area.

Fig 3.11 shows total UL transmit power of all UEs in HetNets area. It is obvious

that JUDUA achieves the lowest total UL transmit power among these three algorithms under the scenarios with different numbers of UEs in HetNets area, which demonstrates that JUDUA has merits in the energy saving.

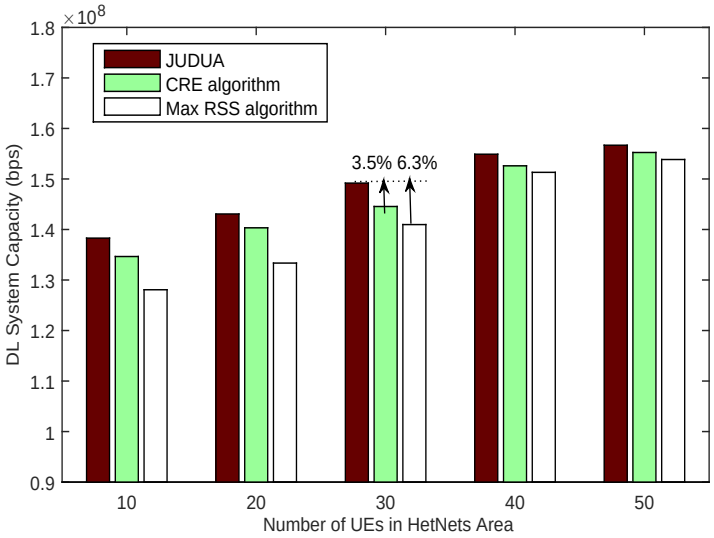


Figure 3.12: DL system capacity versus different numbers of UEs in HetNets area.

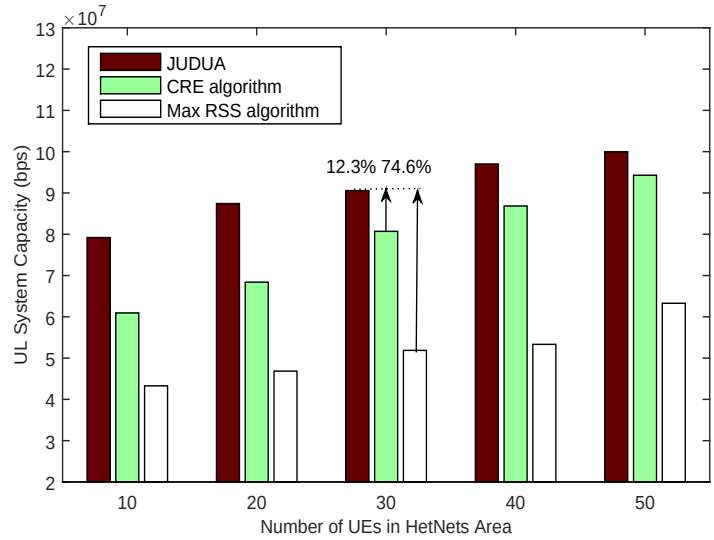


Figure 3.13: UL system capacity versus different numbers of UEs in HetNets area.

Fig 3.12 and Fig 3.13 are the DL and the UL system capacity with different numbers of UEs in HetNets area, respectively. The DL and the UL system capacity plotted here is the sum rate of all users in DL and UL, respectively. These figures demonstrate that

JUDUA outperforms the other two algorithms in terms of the system capacity. Using the scenario with 30 UEs as an example, compared with CRE algorithm and max RSS algorithm, JUDUA improves DL system capacity by 3.5% and 6.5%, respectively, and achieves significant improvement in UL system capacity with 12.3% and 74.6% increase, respectively.

3.3.4.2 Conclusions

Joint UL and DL user association for energy-efficient HetNets, called JUDUA, was proposed. The JUDUA optimisation was modelled as a bargaining problem, which was resolved by NBS. It was shown that for each iteration, JUDUA had a low computational complexity of $O\left(M^2 N \log_2 N + M^3\right)$. Simulations indicate that compared with the conventional user association algorithms, JUDUA provides a significant gain on UL and DL energy efficiencies for most UEs, improving user fairness in terms of the energy efficiency. The merits of JUDUA are further reflected on the decreased UL transmit power, as well as the increased UL and DL system capacity.

3.4 Opportunistic User Association for Multi-Service HetNets Using NBS

In this section, opportunistic user association for multi-service HetNets is proposed for the QoS provision of the delay constraint traffic while providing fair resource allocation for the best effort traffic. Only DL transmission is considered in this algorithm.

3.4.1 Motivation

With the fast development of Internet of Things (IoT) applications, machine-to-machine (M2M) traffic grows explosively. As a great amount of M2M traffic shares the same

network infrastructure with the conventional human-to-human (H2H) traffic, wireless network operators are facing the increasing pressure [SJL+13]. More specifically, it is a big challenge for wireless network operators to accommodate the differentiated QoS requirements of both the M2M traffic and the conventional H2H traffic [ZHW+12].

With the ability to achieve more spectrum-efficient and energy-efficient communications [HM12], HetNets provide a promising architecture for supporting mixed H2H and M2M traffic. The existing research on user association in HetNets laid a solid foundation in understanding user association, nevertheless, the impact of service classification on user association in HetNets is less well understood. This work aims to bridge the gap between user association and QoS support for multi-service traffic in HetNets. With this in mind, an opportunistic user association is proposed to classify the delay constraint traffic as primary service, and the best-effort traffic as secondary service. The proposed opportunistic user association aims at utilising radio resources to support fair resource allocation for secondary service without jeopardising QoS of the primary service.

In line with [LWZY11], there are various types of M2M traffic, such as mobile streaming, smart metering, regular monitoring, emergency alerting and mobile POS (Point Of Sales). Whereas according to the 3GPP standards [3GP02], the H2H traffic can be categorised into four classes: conversational class, streaming class, interactive class and background class. Classifying the traffic into the delay constraint and best effort traffic is able to generally capture the characteristics of all different traffic types, given that the interactive and background class H2H traffic and the regular monitoring and smart metering M2M traffic can be generally mapped to the best-effort traffic, and the others are mapped to the delay constraint traffic.

3.4.2 Problem Formulation

It is assumed there are N UEs in total, where UE_n ($n \in \{1, 2, \cdots, K\}$) requests secondary service (e.g. best effort traffic), denoted as secondary service UE (SSUE), and UE_n

($n \in \{K+1, K+2, \cdots, N\}$) requests primary service (e.g. delay constraint traffic), denoted as primary service UE (PSUE) as shown in Fig. 3.14. The PSUE requires a minimum data rate constraint to ensure the limited delay variation [MHY+12], while the SSUE needs the guaranteed fairness.

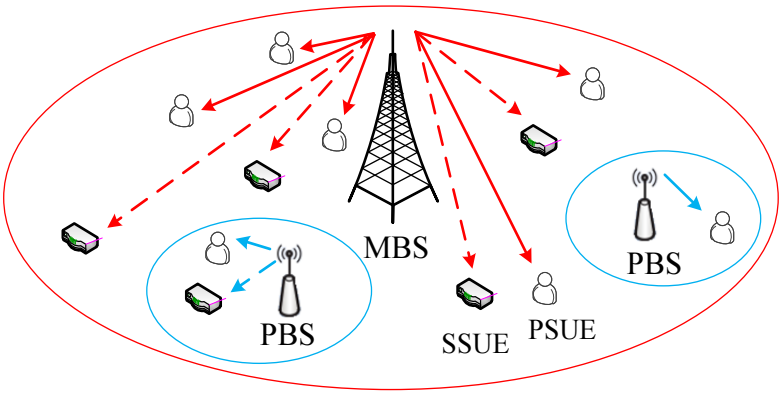


Figure 3.14: System model for opportunistic user association in multi-service HetNets.

In opportunistic user association, BSs are modelled as players in the bargaining problem, and the payoff of BS_m is defined as the sum utility of all the UEs associated with it, which is given by

$$U_m = \sum_{n=1}^{N} x_{mn}\mu_{mn}, \tag{3.32}$$

where u_{mn} is the utility of UE_n associated with BS_m,

$$\mu_{mn} = \begin{cases} b\log\left(r_{mn}^{d}\right), & \forall n \in \{1, 2\cdots, K\} \\ -\exp\left(\frac{-ar_{mn}^{d}}{r_n^{\min}}\right), \forall n \in \{K+1, \cdots, N\} \end{cases}, \tag{3.33}$$

in which a is PSUE's satisfactory factor ($a > 1$), b is SSUE's utility coefficient ($0 < b < 1$) to adjust the utility of PSUE and SSUE into the same order of magnitude, and $r_n^{\min}$ is the required minimum data rate of PSUE.

To encourage user fairness, SSUE's utility is a logarithm function which is concave with diminishing benefits and widely used to construct utility functions [YRC+13].

Fig. 3.15 shows curves of utility of PSUE versus $r_{mn}^d/r_n^{\min}$ with different a. In order to allocate sufficient but not excessive resources to PSUE, in the proposed algorithm, the utility of PSUE is designed to asymptotically approach zero, when the data rate of PUSE is not smaller than the minimum required data rate, that is $r_{mn}^d/r_n^{\min} \geq 1$, otherwise the utility of PSUE is negative. Fig. 3.15 demonstrates if a is very small, say $a = 2$, the utility of PSUE is much smaller than 0, when $r_{mn}^d/r_n^{\min} = 1$. If a is very large, say $a = 50$, the utility of PSUE has already asymptotically approached 0, when $r_{mn}^d/r_n^{\min} < 1$. As such, $a = 10$ [YZZJ11] and $b = 0.005$ are adopted in the simulation.

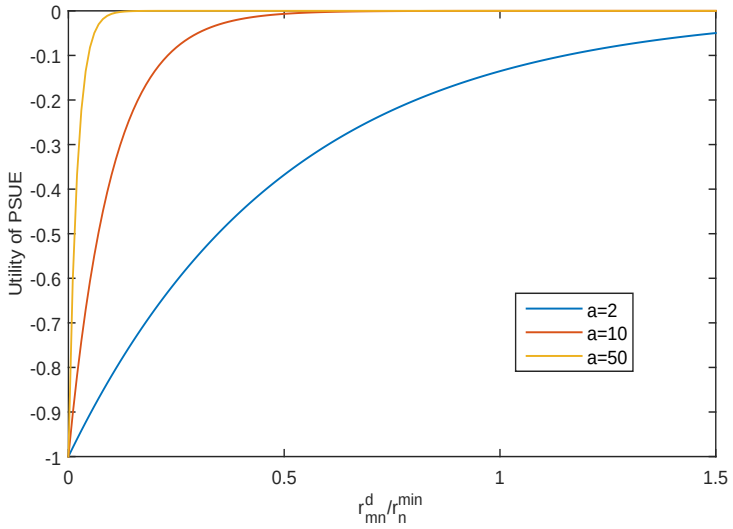


Figure 3.15: Utility of PSUE versus with different values of a.

Substituting (3.33) into (3.32), the payoff of BS_m is

$$U_m = \sum_{n=1}^{K} x_{mn} b \log\left(r_{mn}^d\right) + \sum_{n=K+1}^{N} -x_{mn} \exp\left(\frac{-a r_{mn}^d}{r_n^{\min}}\right). \tag{3.34}$$

Then the optimisation problem of opportunistic user association is formulated as

$$\begin{aligned} \max_{\mathbf{X}} \quad & U = \prod_{m=1}^{M} \left(U_m - U_m^{\min}\right), \\ \text{s.t.} \quad & U_m \geq U_m^{\min}, \ \forall m \\ & x_{mn} = \{0, 1\}, \ \forall m, n \\ & \textstyle\sum_{m=1}^{M} x_{mn} = 1, \ \forall n, \end{aligned} \tag{3.35}$$

where U is the NBS utility, and $U_m^{\min}$ is the minimal payoff of BS_m. Here we set $U_m^{\min} \to 0$, which means PSUE's data rate is larger than the required minimum data rate. $\sum_{m=1}^{M} x_{mn} = 1$ indicates each UE must be associated with a single BS.

The user association problem formulated in (3.35) falls into the class of integer programming problems, where an exact solution usually involves an exhaustive search. However, an exhaustive search is computationally prohibitive when the number of UEs is large. Approximating the problem by its continuous relaxation is a common approach to the integer programming problem. The idea of continuous relaxation is to enlarge the constraint set to include all convex combinations of the original points. As such, $x_{mn} = \{0, 1\}$ is relaxed to $0 \leq x_{mn} \leq 1$, where x_{mn} specifies the probability that UE_n is associated with BS_m.

Note that, although the probabilistic user association is adopted through continuous relaxation here. The user association proposed in Section 3.4.3 determines the optimal deterministic user association. This will be made clear in the proof of **Proposition 3** in Section 3.4.3.

To sum up, the bargaining problem of opportunistic user association in HetNets is described as follows. Each BS_m has the payoff U_m, which is concave and upper-bounded, since its Hessian matrix is negative semidefinite. The optimisation goal is to determine $\mathbf{X}$ to maximise all U_m simultaneously under the constraint $U_m \geq U_m^{\min}$. It is crucial to design a simple and fast method to find the optimal, unique and fair $\mathbf{X}$.

3.4.3 Opportunistic User Association Algorithm

3.4.3.1 Algorithm for Two-BS Case

The same as the Section 3.2.3.1, this subsection focuses on the two-player bargaining algorithm for two BSs ($M = 2$). Then the next subsection proceeds with the multi-player bargaining algorithm for multiple BSs. In the two-player bargaining algorithm, BS_1 and BS_2 are arbitrary parts of BSs in HetNets. They can be modelled as a MBS and a PBS, or two MBSs, or two PBSs. Inspired by the low complexity algorithm in [YC02], the so-called two-band partition is applied to determine the opportunistic user association. It is shown in [YC02] that the two-band partition is near-optimal for the optimisation goal of weighted rate maximisation.

Table 3-F: Two-band UE Partition for Opportunistic User Association

Step 1. Initialisation
Initialise the user association and guarantee BS_m's minimal payoff.
Set $U_{\max}(0) = U$, and calculate A_m, B_m and U_m $m \in \{1, 2\}$.
Set $i = 1$.
Step 2. Sort UEs
Sort UEs from largest to smallest according to $f(\mu_{1n}, \mu_{2n})$.
Step 3. for $n = 1, \cdots, N-1$
Calculate $U(n)$, where UE_1,..., UE_n is associated with BS_1,
and UE_{n+1},..., UE_N is associated with BS_2.
end for
Step 4. Choose the two-band partition $\aleph = \arg\max_n U(n)$
which generates the largest U **satisfying the constraints.**
Set $U_{\max}(i) = U(\aleph)$.
Step 5. Update user association
if $U_{\max}$ cannot be increased by updating A_m, B_m and U_m, the iteration ends;
otherwise, update A_m, B_m and U_m according to the new partition,
set $i = i + 1$, and go to ***step 2***.

The proposed two-band UE partition in opportunistic user association is described in Table 3-F.

Proposition 3: The two-band UE partition shown in Table 3-F is near-optimal to the optimisation problem in (3.35) with $M = 2$.

Proof. In opportunistic user association, the two-player bargaining is to maximise the

NBS utility $U = \left(U_1 - U_1^{\min}\right)\left(U_2 - U_2^{\min}\right)$. Similar to [YC02], the constraint $x_{mn} = \{0, 1\}$ is relaxed to continuous values with $0 \leq x_{mn} \leq 1$. Then the Lagrangian function of (3.35) as a function of x_{mn} is

$$L = \prod_{m=1}^{2}\left(U_m - U_m^{\min}\right) + \sum_{n=1}^{N}\lambda_n\left(\sum_{m=1}^{2}x_{mn} - 1\right), \tag{3.36}$$

where λ_n is the Lagrangian multipliers. By taking the KKT condition and substituting (3.32) into (3.36), the derivative of (3.36) with respect to x_{mn} is

$$\frac{\mu_{1n} + \sum\limits_{n=1}^{N} x_{1n}\frac{d\mu_{1n}}{dx_{1n}}}{U_1 - U_1^{\min}} = \frac{\mu_{2n} + \sum\limits_{n=1}^{N} x_{2n}\frac{d\mu_{2n}}{dx_{2n}}}{U_2 - U_2^{\min}}. \tag{3.37}$$

The left and right side of (3.37) can be interpreted as the marginal benefits of UE_n for BS_1 and BS_2, respectively. When UE_n is only associated with one BS, equation (3.37) becomes inequality. If the left side of (3.37) is greater than the right side, UE_n should be associated with BS_1 and vice versa with BS_2. Substitute (3.34) into (3.37), and take the difference between left and right sides of (3.37), $f\left(\mu_{1n}, \mu_{2n}\right)$ is defined as the difference of marginal benefits of UE_n for BS_1 and BS_2

$$f\left(\mu_{1n}, \mu_{2n}\right) = \frac{\left(\mu_{1n} + A_1 + B_1\right)}{U_1 - U_1^{\min}} - \frac{\left(\mu_{2n} + A_2 + B_2\right)}{U_2 - U_2^{\min}}, \tag{3.38}$$

where

$$A_m = \sum_{n=K+1}^{N}\left(\frac{-ar_{mn}^{d}x_{mn}}{\sum\limits_{n=1}^{N}x_{mn}r_n^{\min}}\right)\exp\left(\frac{-ar_{mn}^{d}}{r_n^{\min}}\right), m \in \{1, 2\}, \tag{3.39}$$

and

$$B_m = \sum_{n=1}^{K} x_{mn}\left(-b \Big/ \sum\nolimits_{n=1}^{N} x_{mn}\right), m \in \{1, 2\}. \tag{3.40}$$

Thus whether UE_n should be associated with BS_1 or BS_2 can be decided by checking whether $f\left(\mu_{1n}, \mu_{2n}\right)$ is greater or less than zero. With fixed A_m, B_m and U_m, the index

of UEs is sorted to make $f(\mu_{1n}, \mu_{2n})$ decrease in n. As such, $f(\mu_{1n}, \mu_{2n})$ is a monotonic function of n. Then (3.38) is similar to the weighted maximisation in [YC02], so the two-band partition is the near-optimal solution.

Within each iteration, U_m is fixed, and the two-band UE partition achieves the near-optimal solution. Then in the next iteration, U_m is updated. Remember U_m^* is the NBS. If $U_1 > U_1^*$ and $U_2 < U_2^*$, $(U_1 - U_1^{\min})^{-1}$ is small and $(U_2 - U_2^{\min})^{-1}$ is relatively large. Consequently, the marginal benefit of BS_1 will be reduced, leading to a disadvantage for bargaining user association in the next iteration, and vice versa. This is one explanation why the proposed two-band UE partition converges to the NBS [HJL05].

It is worthy mentioning that due to the continuous relaxation of x_{mn}, in the end, the UE_n with $f(\mu_{1n}, \mu_{2n}) = 0$ will be associated with BS_1 and BS_2 simultaneously. This UE_n can be named as boundary UE, and all the other UEs are associated with either BS_1 or BS_2, which is similar with the approach in [KBCH10]. With an explanation similar to the one in [YC02], it is noted that given the number of UEs is much larger than the number of BSs, the boundary UE can be associated with either BS arbitrarily without affecting the system performance. □

This two-band UE partition has the complexity of $O\left(N^2\right)$ for each iteration, which can be further improved by the binary search algorithm with a complexity of $O(N\log_2 N)$. According to simulations, this two-band UE partition converges within three rounds.

3.4.3.2 Algorithm for Multi-BS Case with Coalition

The procedure of the algorithm for multi-BS case with coalition is the same as that in Section 3.2.3.2.

3.4.4 Simulation Results and Conclusions

3.4.4.1 Simulation Results

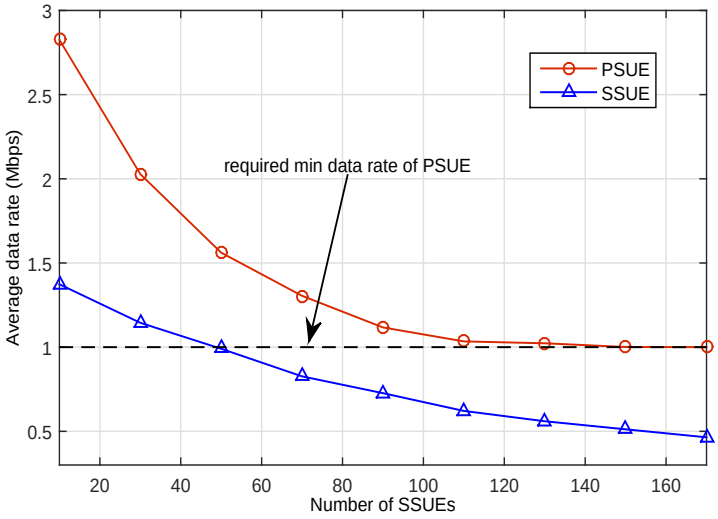


Figure 3.16: Average data rate versus different numbers of SSUEs in HetNets area.

Fig. 3.16 shows the average data rate versus number of SSUEs when 1 MBS and 1 PBS are simulated. Assuming there are 30 PSUEs, the figure indicates that the increase of SSUEs does not affect QoS of PSUEs, since the average data rate of PSUE can always fulfill the minimum data rate requirement.

We then set 20 PSUEs and 80 SSUEs. We define $1 - \exp\left(-10r_n/r_n^{\min}\right)$, $n = K+1, K+2 \cdots, N$ [YZZJ11] to weight the satisfaction degree of PSUEs. The fairness among SSUEs is validated by JFI defined in (3.25).

Fig. 3.17 and Fig. 3.18 are a comparison with the reference algorithm [CFM12] which aims to maximise the sum rate without service classification. Both figures show that in the scenario with more picocells, the proposed opportunistic user association outperforms the reference algorithm in terms of QoS support. The proposed algorithm not only fulfills the minimum data rate requirement of PSUEs, but also improves the fairness among SSUEs.

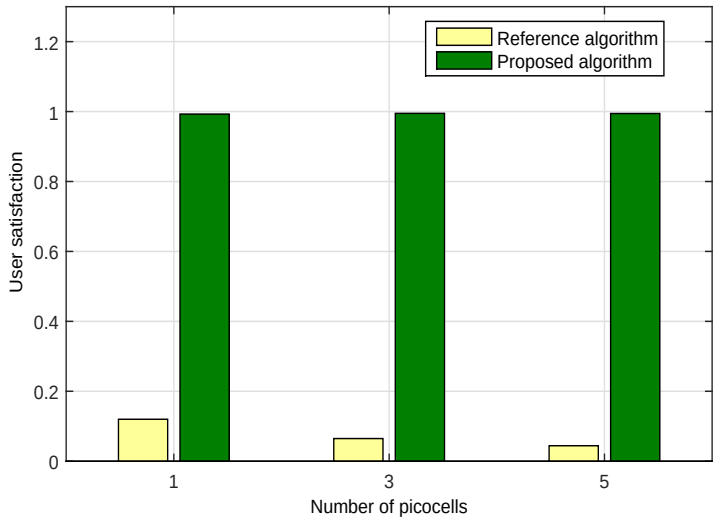


Figure 3.17: Average satisfaction of PSUE versus different numbers of picocells in HetNets area.

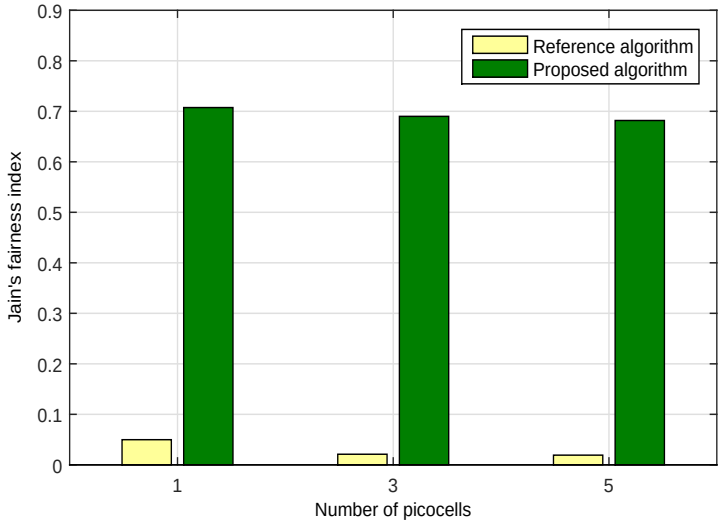


Figure 3.18: Jain's fairness index of SSUE versus different numbers of picocells in HetNets area.

3.4.4.2 Conclusions

An opportunistic user association was proposed for multi-service HetNets, where the delay constraint traffic was classified as primary service and the best effort traffic as secondary service. The opportunistic user association optimisation was modelled as a bargaining problem, which was resolved by NBS. It was shown that for each itera-

tion, the proposed opportunistic user association had a low computational complexity of $O\left(M^2 N \log_2 N + M^3\right)$. Simulations indicate that the proposed algorithm can support fair resource allocation for the best effort traffic without jeopardising QoS of the delay constraint traffic.

3.5 Summary

This chapter focused on the user association optimisation in conventional grid-powered HetNets. More specifically, the user association optimisation was formulated as a bargaining problem from cooperative game theory, and then NBS was applied to achieve the optimal user association solution.

Section 3.1 specified the general system model and simulation platform adopted in this chapter.

The NBS based user association algorithm was developed to improve the DL system performance in Section 3.2. The proposed algorithm has a low computational complexity of $O\left(M^2 N \log_2 N + M^3\right)$ for each iteration. Simulation results validate the merits of the proposed algorithm in effectively offloading users from macrocell to picocells, and improving user fairness, as well as achieving comparable sum rate of all users to the existing max sum rate algorithm.

Through bringing in UL transmission into consideration, a NBS based joint UL and DL user association algorithm, named as JUDUA, was proposed to enhance both UL and DL energy efficiencies in Section 3.3. Compared with the conventional user association algorithms, JUDUA provides a significant gain on UL and DL energy efficiencies for most UEs, and consequently improves user fairness in terms of energy efficiency. In addition, compared with the existing CRE algorithm and max RSS algorithm, JUDUA reduces the UL transmit power by 6.8% and 15.8%, respectively, improves DL system capacity by 3.5% and 6.5%, respectively, and achieves significant improvement in UL

system capacity with 12.3% and 74.6% increase, respectively.

Finally, taking multi-service into consideration, a NBS based opportunistic user association algorithm was proposed for the QoS provision of the delay constraint traffic while providing fair resource allocation for the best effort traffic in Section 3.4. Simulations validate the effectiveness of the proposed algorithm in supporting fair resource allocation for the best effort traffic without jeopardising QoS of the delay constraint traffic.

Chapter 4

User Association Optimisation for HetNets with Renewable Energy Powered BSs

The renewable energy powered BSs via energy harvesting in HetNets are attractive, since they are not only environmentally friendly, but also able to open up entire new categories of low cost *drop and play* deployment, specially of small cells [HSDA14]. This chapter investigates the green HetNets with renewable energy powered BSs. Both the optimal offline and heuristic online algorithms are proposed for the adaptive user association, which are able to adjust the user association decision according to the amount of renewable energy harvested by BSs. Note that this chapter optimises the user association in a snapshot, which amounts to the transmit power of BSs is assumed to be determined by the amount of harvested renewable energy from environment in the problem formulation. The two-dimensional optimisation in both time and space dimensions is extended in Chapter 5.

4.1 Motivation

Integrating energy harvesting capability into BSs entails many challenges to resource allocation algorithm design, due to the randomness of energy availability at renewable energy sources. The HetNets with renewable energy powered BSs call the prior resource alloca-

tion algorithms for conventional grid-powered HetNets into question, as resource allocation algorithms in the renewable energy powered scenario should be adapted according to the energy and load variations across time and space.

In this chapter, the adaptive user association in HetNets with renewable energy powered BSs is investigated, where all BSs are assumed solely powered by the harvested energy from renewable energy sources. BSs across tiers differ in terms of energy harvesting rate, maximum transmit power and deployment density. In conventional grid-powered HetNets, user association is determined based on the assumption that all BSs can transmit with constant powers, whereas transmit powers of BSs vary in HetNets with renewable energy powered BSs , where the transmit power of BS is determined by the amount of harvested energy. The adaptive user association is formulated as an optimisation problem which aims to maximise the number of accepted UEs and minimise the radio resource consumption in the scenario where the available energy of BSs is dependent on the harvested energy in a certain period of time. First an optimal offline algorithm is proposed, where the gradient descent method is used to achieve the pseudo-optimal user association solution. The performance of proposed gradient descent based user association algorithm is verified by simulation results. Considering practical implementation, a heuristic online user association algorithm is further proposed, which is capable of making timely user association decision for incoming UEs based on remaining available network resources. Simulation results indicate the proposed online algorithm achieves a good tradeoff between UEs acceptance ratio and association delay. To the best of my knowledge, this is the first work on user association optimisation in HetNets with renewable energy powered BSs.

4.2 System Model

The 2-tier DL HetNets are considered where tier 1 is modelled as macrocell and tier 2 as picocell. Fig. 4.1 details the system model for adaptive user association in HetNets with

renewable energy powered BSs. The considered HetNets are composed of K macrcocell geographical areas each containing one MBS, denoted as BS_0^k, and N_p PBSs, denoted as BS_m^k ($m \in \{1, \cdots, N_p\}$, $k \in \{1, 2, \cdots, K\}$), and thus there are $K * N_p$ PBSs in HetNets. All BSs share the same frequency band, with frequency reuse factor equaling to one. It is worth mentioning that the proposed user association can be extended to L-tier HetNets ($L > 2$). There are N_u user equipments UE_n ($n \in \{1, 2, \cdots, N_u\}$) in total randomly distributed in HetNets area. It is assumed that at any time each UE can either be associated with a single BS or no BS when the UE is rejected.

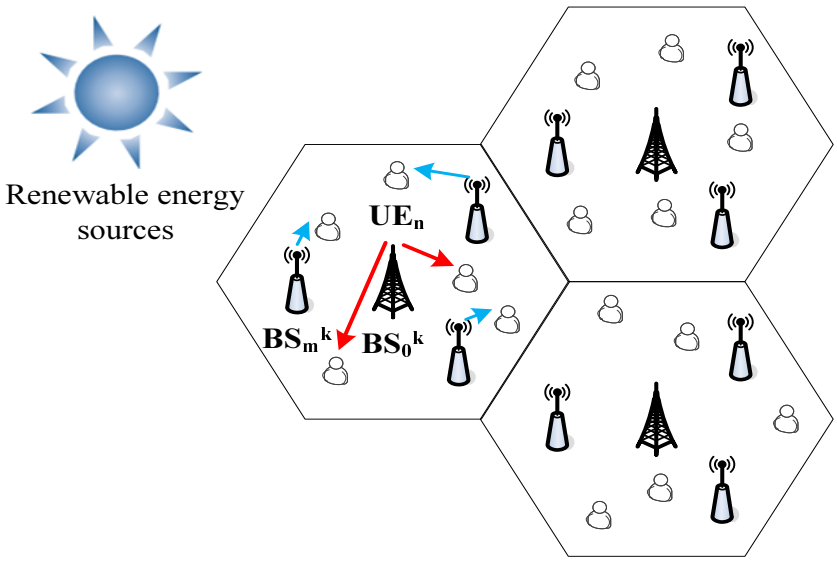


Figure 4.1: System model for adaptive user association in HetNets with renewable energy powered BSs.

In the adaptive user association, each BS is assumed to be solely powered by the harvested energy from renewable energy sources. The BSs from different tiers differ in terms of the energy harvesting rate which determines transmit powers of different BSs. Since the process of energy harvesting is not my research focus, in this chapter the power delivery capacity of energy harvesting for BS_m^k is formulated as Z_m^k, a stationary stochastic process described by probability density function (PDF) $f_m^k\left(z_m^k\right)$ ($m \in \{0, \cdots, N_p\}$) [ZPSY13].

For the sake of tractability, it is assumed that HetNets operate on two time scales [HSDA14]: *i) long time scale*, over which each BS harvests energy from renewable energy sources,

and *ii) short time scale*, over which the user association and scheduling decision are made. Such assumption facilitates analysis since it allows us to assume that each BS transmits to each UE with fixed power in each resource block over short time scale.

To formulate the user association problem, the user association matrix $\mathbf{x} = [x_{mn}^k]$ is defined as

$$x_{mn}^k = \begin{cases} 1, \text{ if } \mathrm{UE_n} \text{ is associated with } \mathrm{BS_m^k} \\ 0, \text{ otherwise, } \mathrm{m} \in \{0, \cdots, \mathrm{N_p}\} \end{cases}. \tag{4.1}$$

The received DL SINR of UE_n when associated with MBS in k-th macrocell area (BS_0^k) is given by

$$\gamma_{0n}^k = \frac{P_0^k\left(\varphi_0^k\right) g_{0n}^k}{\sum\limits_{k' \neq k}^{K} P_0^{k'}\left(\varphi_0^{k'}\right) g_{0n}^{k'} + \sum\limits_{k'=1}^{K} \sum\limits_{m=1}^{N_p} P_m^{k'}\left(\varphi_m^{k'}\right) g_{mn}^{k'} + \sigma_n^2}, \tag{4.2}$$

and the received DL SINR of UE_n when associated with PBS in k-th macrocell area (BS_m^k, $m \in \{1, \cdots, N_p\}$) is as follows

$$\gamma_{mn}^k = \frac{P_m^k\left(\varphi_m^k\right) g_{mn}^k}{\sum\limits_{k'=1}^{K} P_0^{k'}\left(\varphi_0^{k'}\right) g_{0n}^{k'} + \sum\limits_{k'=1}^{K} \sum\limits_{m' \neq m}^{N_p} P_{m'}^{k'}\left(\varphi_{m'}^{k'}\right) g_{m'n}^{k'} + \sigma_n^2}, \tag{4.3}$$

where g_{mn}^k is the average channel power gain between BS_m^k and UE_n, which considers pathloss and shadowing, and σ_n^2 is the estimated noise power level at UE_n. $P_m^k(\cdot)$ ($m \in \{0, \cdots, N_p\}$) is the transmit power of BS_m^k. Since the energy harvesting parameter determines the transmit power of BS, here $P_m^k\left(\varphi_m^k\right)$ is the inverse of $F_m^k\left(z_m^k\right) \triangleq \int_0^{z_m^k} f_m^k(\xi) d\xi$, where φ_m^k is the power outage probability for BS_m^k [ZPSY13].

Based on SINR values, the required radio resources for UE_n when associated with

BS_m^k can be derived as

$$c_{mn}^k = f\left(\gamma_{mn}^k\right) = \frac{\psi_n}{\log_2\left(1+\gamma_{mn}^k\right)}, m \in \{0, 1, \cdots N_p\}, \tag{4.4}$$

where ψ_n is required data rate of UE_n.

4.3 Problem Formulation

The adaptive user association is formulated as an optimisation problem which aims to maximise the number of accepted UEs while minimising the radio resource consumption in the scenario where all BSs in HetNets are solely powered by harvested energy from renewable energy sources.

$$\max G(\mathbf{x}) = (1-w)\sum_{k=1}^{K}\sum_{m=0}^{N_p}\sum_{n=1}^{N_u} x_{mn}^k - w\sum_{k=1}^{K}\sum_{m=0}^{N_p}\sum_{n=1}^{N_u}\rho_m^k c_{mn}^k x_{mn}^k \tag{4.5}$$

s.t.

$$\sum_{n=1}^{N_u} c_{mn}^k x_{mn}^k \leq C_m^k, \quad \forall m \in \{0, \cdots, N_p\} \tag{4.6}$$

$$\sum_{k=1}^{K}\sum_{m=0}^{N_p} x_{mn}^k = 1 \text{ or } 0, \quad \forall n \in \{1, \cdots, N_u\} \tag{4.7}$$

$$x_{mn}^k \in \{0, 1\}, \quad \forall k \in \{1, \cdots, K\}. \tag{4.8}$$

The first term in (4.5) evaluates the number of accepted UEs in HetNets, and the second term evaluates the amount of radio resources used for serving accepted UEs. Thus the maximisation of the objective function (4.5) requires maximisation of the number of accepted UEs in HetNets and minimisation of the radio resource consumption. w specifies the relative importance between the number of accepted UEs and the consumed radio resources. The larger w will lead to more emphasis on the resource utilisation to minimise the resource consumption, and the smaller w will attach more importance to

the number of accepted UEs, in order to improve the network capacity. ρ_m^k is the weight of resources in BS_m^k. Generally the value of weights can be selected as a composite tradeoff among signal quality, spectrum efficiency and load balancing requirements, and they can also be adapted in real time to address the dynamic change in HetNets. Here we set $\rho_m^k = P_m^k\left(\varphi_m^k\right)$, since the BS with larger transmit power may provide higher RSS to users, leading to more associated users, thereby having higher chance to become the early capacity bottleneck. Thus the proposed algorithm with $\rho_m^k = P_m^k\left(\varphi_m^k\right)$ encourages more UEs to be associated with low-power BSs, due to the fact that the resource consumption of the low-power BS has lower weight in the objective function in (4.5). Constraint (4.6) enforces resource constraint at each BS, where C_m^k is the total bandwidth of BS_m^k. Constraint (4.7) indicates that each UE can only be associated with a single BS or no BS at any time.

4.4 Gradient Descent Based User Association Algorithm

The formulated optimisation problem in Section 4.3 is a 0-1 knapsack problem, which is NP-hard. It is difficult to obtain an optimal solution in real time, especially when there are large number of UEs in HetNets. In this section, the gradient descent based user association algorithm is proposed which is an offline algorithm for the pseudo-optimal solution. For a linear optimisation problem, the pseudo-optimal solution approaches the global optimal solution which is located at the boundary of the constraint region [LHWQ12].

In order to apply the gradient descent method, $x_{mn}^k \in \{0,1\}$ is relaxed to $0 \leq x_{mn}^k \leq 1$. In this case, x_{mn}^k indicates the probability that UE_n is associated with BS_m^k. With the aid of the gradient descent method, the value of x_{mn}^k is updated along the direction $\Delta x_{mn}^k = dG\left(\mathbf{x}\right)/dx_{mn}^k$ as

$$x_{mn}^k\left(t\right) = x_{mn}^k\left(t-1\right) + \delta \Delta x_{mn}^k, \tag{4.9}$$

where δ is the step size. The value of x_{mn}^k is updated according to (4.9) until constraint (4.6) reaches the equality. Then x_{mn}^k values are sorted in the descending order. The larger value of x_{mn}^k means higher association probability. UEs are accepted by the HetNets sequentially in the order specified by the sorted x_{mn}^k values. Each UE can only be accepted by one BS at any time. Once a UE is accepted, the constraints (4.6) and (4.7) are checked. The whole procedures stop when all UEs are accepted by HetNets, or all constraints are reached. The detailed gradient descent based user association algorithm can be described in Algorithm 1.

Algorithm 1: Gradient Descent Based User Association Algorithm in HetNets with Renewable Energy Powered BSs

Step 1. Initialisation
set $x_{mn}^k = 0$, $acceptUEset = \emptyset$, $acceptUEnum = 0$,
$\mathbf{BSrest}^k = [C_0^k, \cdots C_{N_p}^k]$, $\Delta x_{mn}^k = dG(\mathbf{x})/dx_{mn}^k$

Step 2. Update the x_{mn}^k values
for $k = 1, \cdots K$
  for $m = 0, \cdots N_p$
  $usedBW_m^k = \sum_{n=1}^{N_u} x_{mn}^k c_{mn}^k$
    while $(usedBW_m^k < C_m^k)$
      for $n = 1, \cdots N_u$
      $x_{mn}^k = x_{mn}^k + \delta \Delta x_{mn}^k$
      end for
    $usedBW_m^k = \sum_{n=1}^{N_u} x_{mn}^k c_{mn}^k$
    end while
  end for
end for

Step 3. User association according to x_{mn}^k values
$[\mathbf{K}, \mathbf{M}, \mathbf{N}] = \text{sort}\left(x_{mn}^k, \text{descent}\right)$
$i = 0$
while $(acceptUEnum < totalUEnum)$ & $\left(\mathbf{BSrest}^k > \mathbf{0}\right)$
  $i = i + 1$
  if $(N(i) \notin acceptUEset)$ & $\left(BSrest_{M(i)}^{K(i)} - c_{M(i)N(i)}^{K(i)} \geq 0\right)$
  $acceptUEnum = acceptUEnum + 1$
  $acceptUEset = acceptUEset \cup N(i)$
  $BSrest_{M(i)}^{K(i)} = BSrest_{M(i)}^{K(i)} - c_{M(i)N(i)}^{K(i)}$
  end if
end while

Given that

$$\Delta x_{mn}^k = dG(\mathbf{x}) \Big/ dx_{mn}^k = (1 - w) - w\rho_m^k c_{mn}^k, \tag{4.10}$$

and UEs which want to be associated with the same BS have the same ρ_m^k, UE_n which requires less radio resource from BS_m^k has higher association probability. As for different

BSs, the less ρ_m^k is, the comparatively higher the association probability with BS_m^k is. Consequently, the proposed gradient descent based user association algorithm encourages more UEs to be associated with the low-power PBSs, thereby enhancing the load balancing throughout HetNets.

4.5 Heuristic Online User Association Algorithm

In reality, UEs arrive and depart dynamically and quick user association decision should be made. It is not possible to wait until all requesting UEs arrive, and then make the global optimal user association as the optimal offline algorithm, which achieves higher network capacity at the cost of longer processing time and larger user association delay. In real networks, since HetNets have already accepted some UEs which have occupied some radio resources, the user association decision for incoming UEs can only be made based on remaining radio resources in HetNets.

Inspired by the gradient descent based user association algorithm described in Section 4.4, where UEs will associate with the BS from which they require the least weighted radio resources, in the section, a heuristic online user association algorithm is proposed in HetNets with renewable energy powered BSs as described in Algorithm 2. The heuristic online user association algorithm is implemented over the *short time scale*, where the transmit power of BS is assumed as constant and pre-determined by the energy harvesting parameters. The heuristic online user association algorithm is activated whenever a UE attempts the network access or is handovered to another BS over the *short time scale*. The proposed online algorithm associates the new arrival UE with the best serving BS based on the availability of remaining radio resources, and UEs which have already been accepted by a certain BS will not be redirected to another BS.

The Φ_m described in Algorithm 2 indicates weighted bandwidth consumption when UE_n is associated with BS_m^k, which considers the load condition of BS_m^k. The weight is set as $\rho_m^k = P_m^k\left(\varphi_m^k\right)$, in order to associate more UEs with low-power PBSs and achieve

Algorithm 2: Heuristic Online User Association Algorithm in HetNets with Renewable Energy Powered BSs

Step 1. Initialisation
set $\rho_m^k = P_m^k\left(\varphi_m^k\right)$, $BSrest_m^k(0) = C_m^k$,
$usedBW_m^k(0) = 0$, $k \in \{1, \cdots, K\}$, $m \in \{0, \cdots, N_p\}$

Step 2. Online user association
whenever any UE_n tries to access network
for $k = 1, \cdots K$
  for $m \in \{0, \cdots, N_p\}$
    if $BSrest_m^k(t) - c_{mn}^k \geq 0$
    $\Phi_m^k = usedBW_m^k(t) \cdot \rho_m^k + c_{mn}^k$
    else
    $\Phi_m^k = +\infty$
    end if
  end for
end for
$[k*, m*] = \arg \min_{m \in \{0, \cdots N_p\}, k \in \{1, \cdots K\}} \left(\Phi_m^k\right)$ and $\Phi* = \min_{m \in \{0, \cdots N_p\}, k \in \{1, \cdots K\}} \left(\Phi_m^k\right)$
if $\Phi* \neq +\infty$
UE_n is associated with BS_{m*}^{k*}.
$BSrest_{m*}^{k*}(t+1) = BSrest_{m*}^{k*}(t) - c_{m*n}^{k*}$
$usedBW_{m*}^{k*}(t+1) = usedBW_{m*}^{k*}(t) + c_{m*n}^{k*}$
else
UE_n is rejected.
end if

load balancing throughout HetNets. The weighted bandwidth consumption is used in ranking BSs during the user association. UE will choose the BS with the least weighted bandwidth consumption. The heuristic online algorithm intends to jointly consider the signal quality and load distribution, which will maximise the spectral efficiency and system capacity.

4.6 Simulation Platform and Results

4.6.1 Simulation Platform

Fig. 4.2 illustrates the flowchart of the simulation platform used to evaluate proposed algorithms in HetNets with renewable energy powered BSs. The functionality of each module is summarised as follows.

A. Network initialisation

This module initialises the network topology, the user distribution and system parame-

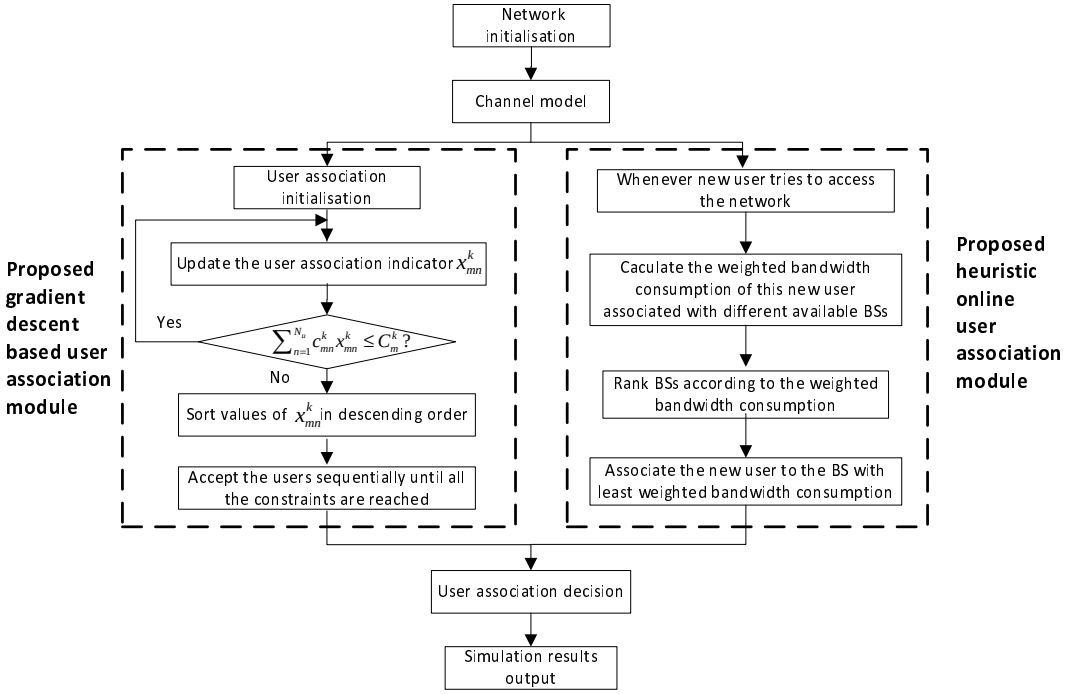


Figure 4.2: Flowchart of simulation platform for user association optimisation in HetNets with renewable energy powered BSs.

ters, such as the harvested power distribution, etc.

- Network topology

  To evaluate the performance of proposed algorithms, the 2-tier DL HetNets are simulated. The theoretical analysis throughout this chapter is independent with the spatial distribution of BSs. In the simulation, locations of all BSs are modelled to be fixed. The simulated HetNets are composed of 19 macrocells. In each macrocell, 3 PBSs are symmetrically located along a circle with radius 200m and MBS in the centre. The inter-site distance is 500m. Fig. 4.3 illustrates the simulated network topology.

- User distribution

  In each snap shot of simulation, users are randomly distributed in HetNets geographically area.

- Harvested power distribution

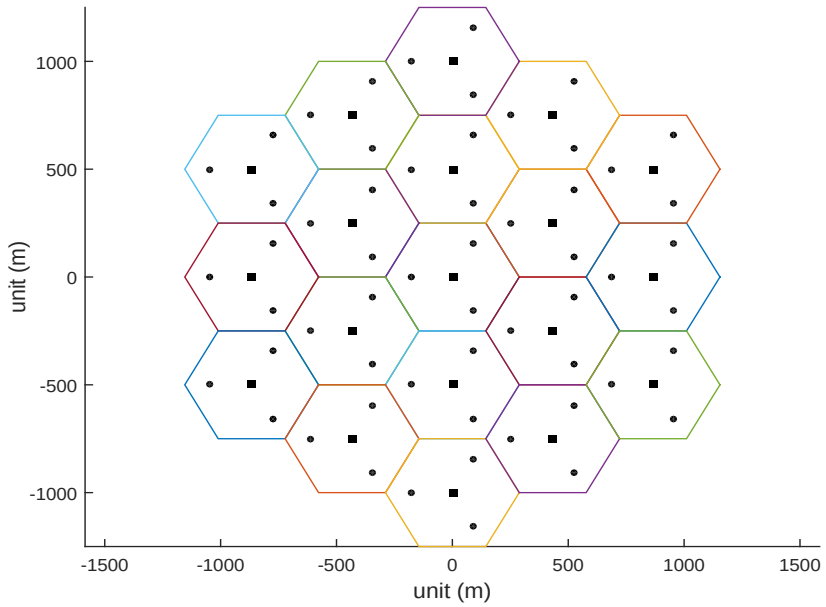


Figure 4.3: Simulated network topology, where MBSs (black square) are located in the centre of the hexagonal cell and PBSs (black circle) are located along the MBSs.

As the distribution of harvested power for BS_m^k is scenario-dependent, for the sake of simplicity, the harvested power is assumed to be uniformly distributed [ZPSY13], i.e. $f_m^k\left(z_m^k\right) = 1/\left(b_m^k - a_m^k\right),\ \forall z_m^k \in [a_m^k, b_m^k]$, with a_m^k and b_m^k as the minimum and the maximum harvested power of BS_m^k, respectively. Note that proposed algorithms can be applied to any other harvested power distribution scenarios.

B. Channel model

This module specifies the channel model in the simulation scenario, and then calculates the channel power gain according to users' and BSs' physical locations. The same channel model is employed here as the model in Section 3.1.2. The details are summarised in Table 4-A.

C. Proposed gradient descent based user association

This module implements the proposed gradient descent based user association algorithm.

D. Proposed heuristic online user association

This module implements the proposed heuristic online user association algorithm.

The other simulation parameters are shown in Table 4-A.

Table 4-A: Simulation Parameters

Parameter	**Value**
Bandwidth	10 MHz [3GP10]
Inter site distance	500 m [3GP10]
Min harvested power of MBS	[0, 20] dBm
Max harvested power of MBS	[20, 50] dBm
Min harvested power of PBS	[0, 10] dBm
Max harvested power of PBS	[10, 34] dBm
Noise power density	-174 dBm/Hz
Pathloss between MBS to UE	$128.1 + 37.6\log_{10}d\,(km)$ [3GP10]
Pathloss between PBS to UE	$140.7 + 36.7\log_{10}d\,(km)$ [3GP10]
Log-normal shadowing fading SD	10 dB [3GP10]
User required data rate	100 Kbps [HJL05]
Power outage probability	5% [ZPSY13]
Number of drops	500

4.6.2 Simulation Results

The proposed gradient descent based user association algorithm (gradient descent algorithm) is compared with the conventional max RSS algorithm.

Fig. 4.4 shows ratios of accepted UEs in the proposed gradient descent algorithm with various values of ρ_m^k and w (where ρ_m^k is the weight of resources in BS_m^k, and w specifies the relative importance between the number of accepted UEs and the consumed radio resources). The ratio of accepted UEs when $w = 10^{-20}$ is larger than that when $w = 10^{-10}$, which validates the design of the proposed algorithm, that the smaller w attaches more importance on the network capacity, leading to the higher UEs acceptance ratio. The figure also verifies that the performance of the proposed gradient descent algorithm when $\rho_m^k = P_m^k\left(\varphi_m^k\right)$ is better than that when $\rho_m^k = 1$, due to the fact that $\rho_m^k = P_m^k\left(\varphi_m^k\right)$ increases the weight of radio resource consumption at high-power BSs, and thus encourages more UEs to be associated with low-power BSs, resulting in the load balancing and higher UEs acceptance ratio throughout HetNets.

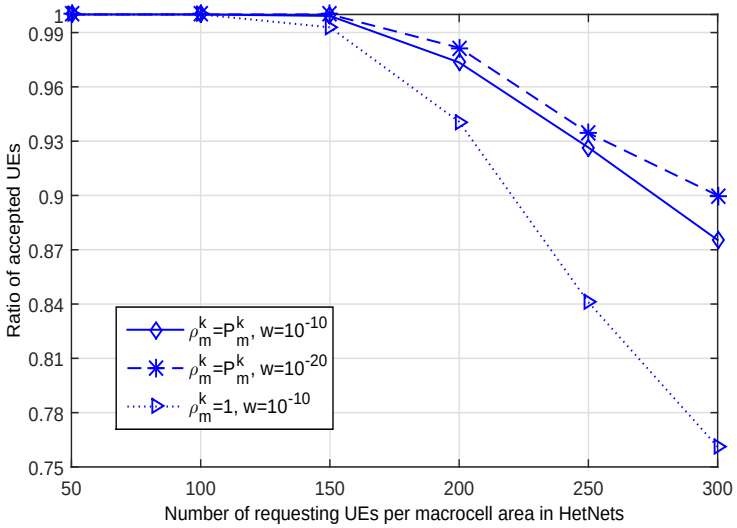


Figure 4.4: Ratio of accepted UEs of the proposed gradient descent algorithm with various values of ρ_m^k and w.

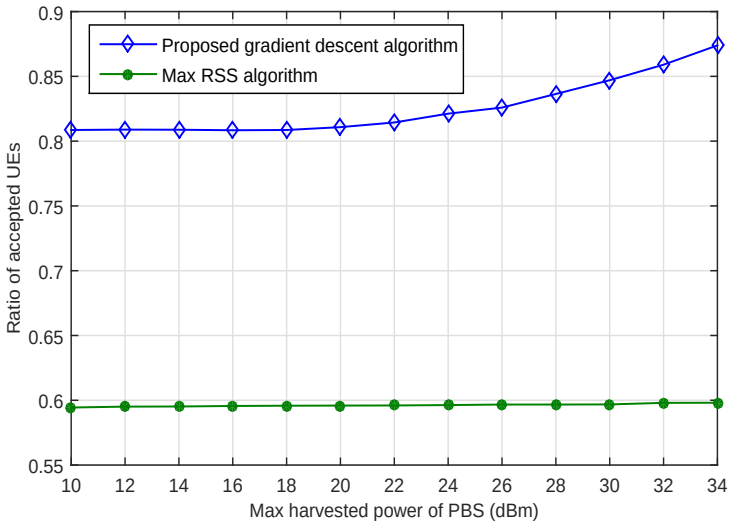


Figure 4.5: Ratio of accepted UEs versus maximum harvested power of PBS.

Then we set $w = 10^{-10}$, $\rho_m^k = P_m^k\left(\varphi_m^k\right)$ and 300 UEs in each marcocell area. It is assumed that the harvested power or transmit power of MBS is 46dBm, and the minimum harvested power of PBS is 5dBm. Fig. 4.5 is ratios of accepted UEs in two algorithms with different maximum harvested powers of PBS. The figure illustrates that in the proposed gradient descent algorithm, the ratio of accepted UEs grows with the increase of maximum harvested power of PBS. The reason is that the larger harvested power will lead to higher transmit power, and enabled by the load balancing property of

the proposed algorithm, more UEs can be accepted. In contrast, the ratio of accepted UEs in the max RSS algorithm does not improve that much. This is because regardless of the value of PBS' maximum harvested power, the transmit power of MBS always dominates the transmit power of PBS, leading to associating the most of UEs with the MBS, and thus the improvement of the maximum harvested power of PBS does not enhance the ratio of accepted UEs. It is also observed that the proposed gradient descent algorithm converges quite quick, and does not incur much more computational complexity, compared with the max RSS algorithm.

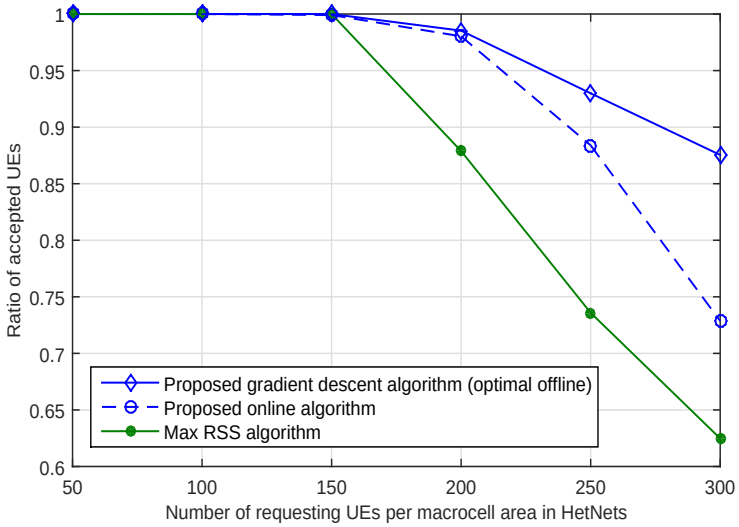


Figure 4.6: Ratio of accepted UEs versus different numbers of requesting UEs.

We then still set $w = 10^{-10}$, $\rho_m^k = P_m^k\left(\varphi_m^k\right)$, and Fig. 4.6 shows the comparison of the accepted UEs ratios in the three algorithms with increasing number of requesting UEs in each marcocell area. The figure indicates that the increase of requesting UEs will decrease the ratio of accepted UEs, but the proposed gradient descent algorithm is superior to the max RSS algorithm. This is because the proposed algorithm takes the resource consumption into account when deciding the user association. As such, the resource is better utilised and more UEs can be served.

The figure also indicates when there are more than 150 requesting UEs, the UEs acceptance ratio in the proposed online algorithm is inferior to that in the proposed optimal offline algorithm. This is because the optimal offline algorithm can obtain the

global optimal result, while the online algorithm can only achieve local optimal, since the user association decision for incoming UEs can only be made based on remaining radio resources in HetNets. When there are 300 requesting UEs, the proposed online algorithm loses 14% UEs acceptance ratio compared with the proposed optimal offline algorithm, but the proposed online algorithm achieves much lower processing delay and shorter waiting time for UEs. Thus the proposed online algorithm achieves a good tradeoff between UEs acceptance ratio and association delay.

4.7 Summary

In this chapter, adaptive user association was investigated in HetNets with renewable energy powered BSs. First an optimal offline algorithm was proposed, the advantages of which were verified by simulation results. For the sake of real-time user association, a heuristic online user association algorithm was further developed. Simulation result indicates the proposed online algorithm achieves a good tradeoff between UEs acceptance ratio and association delay. The proposed algorithms provide insights on how BSs and UEs should associate in HetNets with renewable energy powered BSs.

Chapter 5

User Association Optimisation for HetNets with Hybrid Energy Sources

As mentioned in Chapter 2, BSs powered by hybrid energy sources are preferable over those solely powered by renewable energy sources in order to support uninterrupted service [NLS13, HA12]. As such, this chapter investigates the promising HetNets scenario, where BSs are envisioned to be powered by both power grid and renewable energy sources. For the user association designed for such networks, the vital issue is to minimise the on-grid energy consumption as well as guarantee the user QoS provision. The existing works [HA12, WKLY15, HA13] mentioned in Section 2.3.2 consider the QoS provision in terms of either minimal SINR as in [HA12, HA13] or data rate requirement as in [WKLY15]. Different with them, this chapter incorporates both the requited traffic amount and average traffic delay to support the QoS provision for users. In addition, distinct with the existing research for hybrid energy powered networks which focuses on the minimisation of total on-grid energy consumption only, both total and peak on-grid energy consumption minimisations are considered in this chapter to further reduce OPEX and maintain profitability for mobile network operators. This chapter is organised as follows. The general system model and simulation platform employed in this chapter are introduced in Section 5.1. An optimal user association algorithm is developed to achieve the tradeoffs between average traffic delay and on-grid energy consumption in

Section 5.2. Then the two-dimensional optimisation on user association and green energy allocation is carried out to minimise both total and peak on-grid energy consumptions, as well as enhance the QoS provision in Section 5.3.

5.1 System Model and Simulation Platform

This section elaborates the system model and simulation platform employed in this chapter.

5.1.1 System Model

The 2-tier DL HetNets are considered as shown in Fig. 5.1, where tier 1 is modelled as macrocell and tier 2 as picocell. MBSs provide basic coverage, whereas PBSs are deployed in the coverage area of each MBS to enhance capacity. In this model, a macrocell geographical area $\mathcal{L} \subset \mathbb{R}^2$ is served by a set of BSs $\mathcal{B}$ including one MBS and several PBSs, where all BSs are assumed to share the same frequency band. Let $x \in \mathcal{L}$ denote a location and $i \in \mathcal{B}$ index i-th BS, where $i = 1$ indicates the MBS and the others are PBSs. Assuming the duration of time is divided into $|\mathcal{T}|$ time slots, the length of each time slot is τ seconds and $t \in \mathcal{T}$ denotes the t-th time slot. All BSs in this model are powered by both power grid and renewable energy sources. In order to provide a general model, no particular type of renewable energy source is assumed here. Then the proposed user association algorithm in Section 5.2 is applied to decide which BS within a macrocell geographical area $\mathcal{L}$ will serve which user at location x. The proposed algorithm for the two-dimensional optimisation in Section 5.3 is carried out within $\mathcal{L}$ and across all time slots $\mathcal{T}$ in order to optimise the on-grid energy consumption.

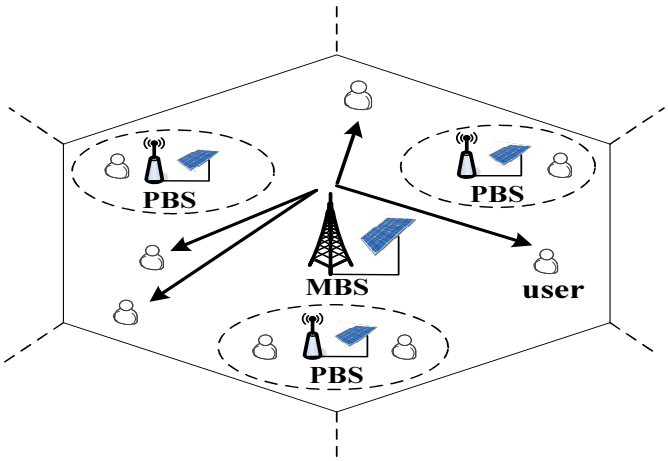


Figure 5.1: System model for user association optimisation in HetNets with hybrid energy sources.

5.1.1.1 Traffic Model

In the presentation of this subsection, the time slot index t is omitted for simplicity. It is assumed that traffic requests arrive according to a inhomogeneous Poisson point process with the arrival rate per unit area $\lambda(x)$, and the traffic size is independently distributed with mean $\mu(x)$. Here $\lambda(x)\mu(x)$ captures the spatial traffic variability, where a hot spot can be characterised by a higher arrival rate or a larger file size.

Assuming a mobile user at location x is associated with BS i, the transmission rate to this user $r_i(x)$ can be generally expressed according to Shannon Hartley theorem [KdVYV12], with W denoted as the operating bandwidth,

$$r_i(x) = W\log_2\left(1 + \mathrm{SINR}_i(x)\right), \tag{5.1}$$

where

$$\mathrm{SINR}_i(x) = \frac{p_i g_i(x)}{\sum_{k\in\mathcal{B},k\neq i} p_k g_k(x) + \sigma^2}, \tag{5.2}$$

here p_i is the transmit power of BS i and σ^2 is the noise power level. The channel power gain $g_i(x)$ between BS i and the user at location x includes pathloss and shadowing. Note that fast fading is not considered here since the time scale of user association is

much larger than the time scale of fast fading. As such, $r_i(x)$ can be considered as a time-averaged transmission rate [SKYK11].

In order to guarantee the QoS of users in a sense that all users are served with the required traffic amount, the fraction of resource blocks allocated by BS i to the user at location x is derived as

$$\varrho_i(x) = \frac{\lambda(x)\,\mu(x)\,y_i(x)}{r_i(x)}, \tag{5.3}$$

where $y_i(x)$ is the user association indicator, if user at location x is associated with BS i, $y_i(x) = 1$, otherwise $y_i(x) = 0$. It is assumed that at each time, one user can only associate with a single BS, and thus $\sum_{i\in\mathcal{B}} y_i(x) = 1$. $\varrho_i(x)$ denotes the average traffic load density of BS i.

Based on the traffic load density, the set $\mathcal{F}$ of feasible loads of BSs $\boldsymbol{\rho} = \left(\rho_1, \ldots, \rho_{|\mathcal{B}|}\right)$ is given by

$$\begin{aligned} \mathcal{F} = \{\boldsymbol{\rho}|\ & \rho_i = \textstyle\int_{\mathcal{L}} \varrho_i dx, \\ & 0 \le \rho_i \le 1-\varepsilon, \textstyle\sum_{i\in\mathcal{B}} y_i(x) = 1, \\ & y_i(x) \in \{0,1\}, \forall x, \forall i\}, \end{aligned} \tag{5.4}$$

where ε is an arbitrarily small positive constant to ensure $\rho_i < 1$.

5.1.1.2 Energy Consumption Model

It is assumed that both MBSs and PBSs in HetNets are powered by hybrid energy sources: power grid and renewable energy sources. Each BS is capable of harvesting green energy from renewable energy sources. Note that small BSs such as PBSs may have smaller energy harvesting rates than those of MBSs. It is also assumed there is no energy transfer among BSs. Each BS will allocate green energy harvested by itself in each time slot. If the allocated green energy is not sufficient to support traffic load, the

BS will consume the energy from the power grid.

Generally, BSs consist of two types of energy consumption: static energy consumption and adaptive energy consumption. Adaptive energy consumption is related to transmit powers of BSs and is typically linear to loads of BSs [AGD+11]. Static energy consumption is the energy consumption when BS is idle without any traffic load. Here the linear approximation of BS energy consumption model in [AGD+11] is adopted, with $E_i(t)$ denoted as the energy consumption of BS i at t-th time slot,

$$E_i(t) = \Delta_i p_i \rho_i(t) \tau + E_i^s, \tag{5.5}$$

where Δ_i is the slope of load-dependent energy consumption of BS i, p_i is transmit power of BS i, τ is the length of each time slot, $\rho_i(t)$ is the traffic load of BS i at t-th time slot, and E_i^s is the static energy consumption of BS i in each time slot. It is worthwhile mentioning that the small BSs such as PBSs and femto BSs generally have smaller static energy consumptions than those of MBSs since they have neither big power amplifiers nor cooling equipments.

Then the green energy allocation of BS i at t-th time slot is denoted as $G_i(t)$, and the on-grid energy consumption of BS i at t-th time slot is expressed as

$$E_i^{grid}(t) = \max\left(E_i(t) - G_i(t), 0\right). \tag{5.6}$$

5.1.2 Simulation Platform

Fig. 5.2 illustrates the flowchart of the simulation platform used in this chapter. The functionality of each module is summarised as follows.

A. Network initialisation

This module initialises the network topology, the user distribution, as well as system parameters.

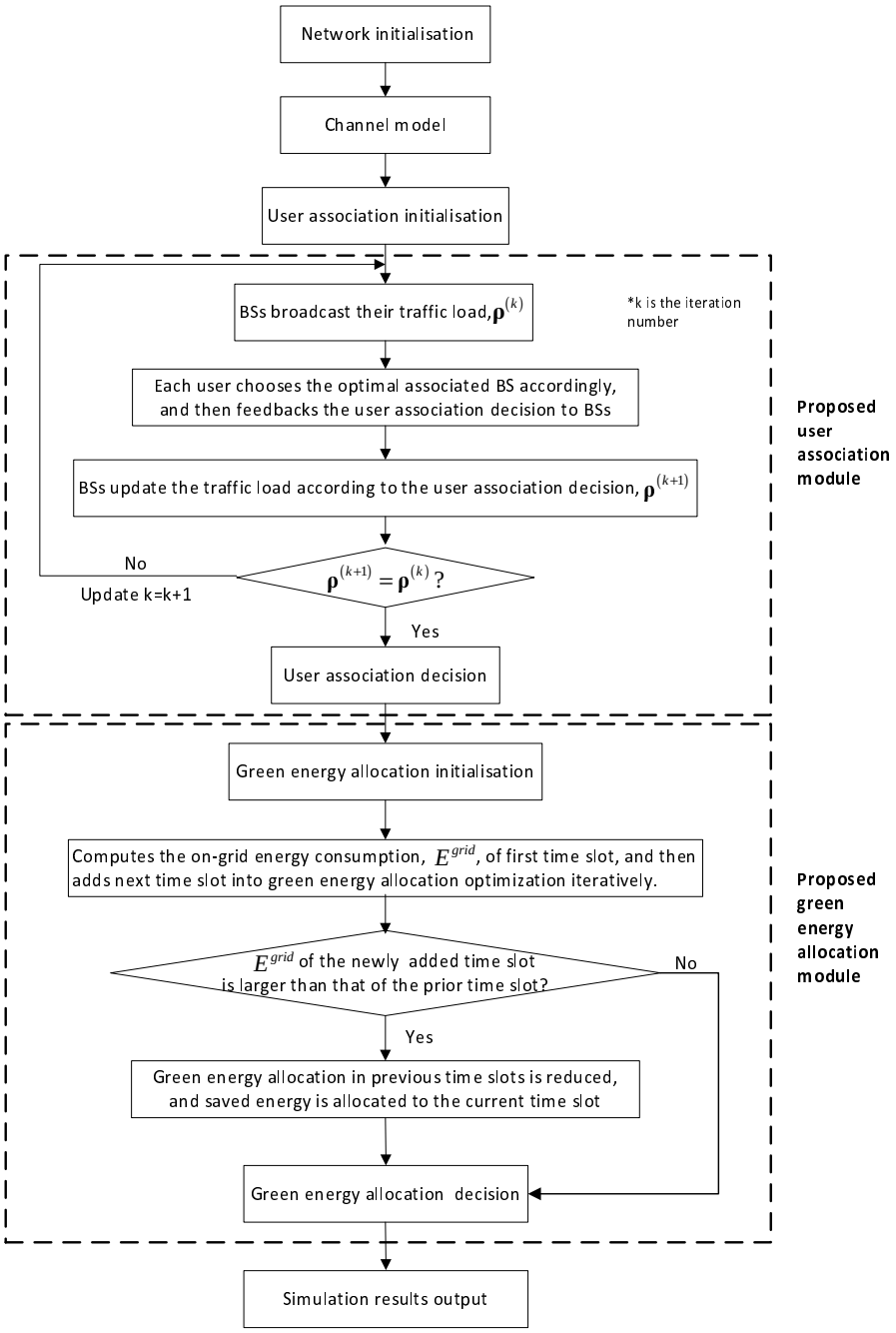


Figure 5.2: Flowchart of simulation platform for user association optimisation in HetNets with hybrid energy sources.

- Network topology

  To evaluate the performance of proposed algorithms, the 2-tier DL HetNets are simulated. The theoretical analysis throughout this chapter is independent with the spatial distribution of BSs. In the simulation, locations of all BSs are modelled to

be fixed. The simulated HetNets are composed of 19 macrocells. In each macrocell, 3 PBSs are symmetrically located along a circle with radius 200m and MBS in the centre. The inter-site distance is 500m. The simulated network topology is the same as Fig. 4.3 in Section 4.6 .

- Traffic model

  For the sake of simplicity, the file transfer requests are simulated to follow a homogenous Poisson point process where $\lambda(x) = \lambda$ [KdVYV12], but note that this model can still apply to the scenario with heterogeneous traffic distribution. Each request is assumed to have exactly one file with mean file size μ as 100 Kbits.

B. Channel model

This module specifies the channel model in the simulation scenario, and then calculates the channel power gain according to users' and BSs' physical locations. The same channel model is employed here as the model in Section 3.1.2. The details are summarised in Table 5-A.

C. Proposed user association module

This module implements the user association algorithm proposed in Section 5.2.3 and Section 5.3.3.1, which will be used in the performance evaluation in Section 5.2.5 and Section 5.3.5.

D. Proposed green energy allocation module

This module implements the optimal offline green energy allocation algorithm proposed in Section 5.3.3.2, which will be used in the performance evaluation in Section 5.3.5.

5.1.2.1 Simulation Parameters

The other simulation parameters are shown in Table 5-A.

Table 5-A: Simulation Parameters

Parameter	**Value**
Bandwidth	10 MHz
Inter site distance	500 m
Transmit power MBS	46 dBm
Transmit power PBS	30 dBm
Noise power density	-174 dBm/Hz
Pathloss between MBS and user	$128.1 + 37.6\log_{10}d\,(km)$ [3GP10]
Pathloss between PBS and user	$140.7 + 36.7\log_{10}d\,(km)$ [3GP10]
Log-normal shadowing fading SD	10 dB [3GP10]
Static power consumption of MBS	780 W [AGD$^+$11]
Static power consumption of PBS	13.6 W [AGD$^+$11]
Slope of MBS load-dependent energy consumption	4.7 [AGD$^+$11]
Slope of PBS load-dependent energy consumption	4.0 [AGD$^+$11]
Number of drops	500

5.2 Optimal User Association for Delay-Energy Tradeoffs in HetNets with Hybrid Energy Sources

Section 5.2 focuses on the user association optimisation in one time slot, and the time slot index t is omitted in the presentation in this section. It is assumed that the green energy generation is constant during one time slot. Furthermore, the harvested green energy cannot be stored, which means if the green energy cannot be consumed once being harvested, it will be wasted.

5.2.1 Motivation

As mentioned in Section 2.3.2, an earlier line of research studies the coordinated MIMO systems [CLW12], power allocation [GZN13] and network planning [ZPSY13] in the hybrid energy sources powered scenario. Different from the existing works, this section focuses on user association optimisation in HetNets with hybrid energy sources. The power grid operation is costly and non-environmentally friendly. In contrast, the harvested energy is green, sustainable and free of cost. Thus instead of the overall network energy consumption minimisation, the objective is to reduce on-grid energy consumption

by maximising the utilisation of green energy harvested from renewable energy sources.

On the other hand, in wireless networks, energy saving is often achieved at the price of degradation in network QoS (i.e., higher latency and lower throughput) [NLW12]. Thus in this section, the proposed user association algorithm aims to achieve the optimal tradeoff between average traffic delay and on-grid energy consumption. To this end, a convex optimisation problem is formulated to minimise the weighted sum of cost of average traffic delay and cost of on-grid energy consumption. It has been proven that the proposed user association algorithm converges to the global optimum.

5.2.2 Problem Formulation

Since $y_i(x) \in \{0,1\}$, $\mathcal{F}$ in (5.4) is not a convex set. In order to facilitate the convex optimisation problem formulation, $y_i(x) \in \{0,1\}$ is relaxed to $0 \leq y_i(x) \leq 1$, where $y_i(x)$ specifies the probability that the user at location x is associated with BS i. Then the updated set $\tilde{\mathcal{F}}$ of feasible loads of BSs $\boldsymbol{\rho} = \left(\rho_1, \ldots, \rho_{|\mathcal{B}|}\right)$ is

$$\begin{aligned} \tilde{\mathcal{F}} = \{\boldsymbol{\rho} |\ & \rho_i = \textstyle\int_{\mathcal{L}} \varrho_i dx, \\ & 0 \leq \rho_i \leq 1-\varepsilon, \textstyle\sum_{i\in\mathcal{B}} y_i(x) = 1, \\ & 0 \leq y_i(x) \leq 1, \forall x, \forall i\}. \end{aligned} \tag{5.7}$$

The authors in [KdVYV12] have proven that the set $\tilde{\mathcal{F}}$ is convex.

Then the user association is formulated as a convex optimisation problem which aims to reduce on-grid energy consumption by optimising the utilisation of green energy harvested from renewable energy sources, as well as enhance network QoS by minimising the average traffic delay of all BSs. The problem is to find optimal load of BSs $\boldsymbol{\rho}$ that minimises the total system cost which is given by

$$\min_{\boldsymbol{\rho}} \left\{ f(\boldsymbol{\rho}) = \varphi(\boldsymbol{\rho}) + \omega\phi(\boldsymbol{\rho}) \,|\boldsymbol{\rho} \in \tilde{\mathcal{F}} \right\}, \tag{5.8}$$

where $\varphi(\boldsymbol{\rho})$ is the cost of average traffic delay and $\phi(\boldsymbol{\rho})$ is the cost of on-grid energy consumption, which will be detailed in the following subsection. $\omega \geq 0$ is the relative weight to balance the tradeoff between average traffic delay and on-grid energy consumption.

Remark 1: Although the user association optimisation problem is formulated via probabilistic user association $\tilde{\mathcal{F}}$, the proposed user association algorithm in Section 5.2.3 determines the optimal deterministic user association. This will be made clear in the proof of **Theorem 1** and **Theorem 2** in Section 5.2.3.

5.2.2.1 Cost Function of Average Traffic Delay

The cost function of average traffic delay is defined as

$$\varphi(\boldsymbol{\rho}) = \sum\nolimits_{i\in\mathcal{B}} \frac{\rho_i}{1-\rho_i}. \tag{5.9}$$

Users associated with the same BS are assumed to be served on the round robin fashion. Considering a dynamic system where new users (traffic requests) arrive randomly (e.g, Poisson process) into the system and leave after being served, the dynamics of this system can be captured by the M/GI/1 multi-class processor sharing system in [J.W98]. In this context, the average number of flows at BS i can be presented as $\rho_i/(1-\rho_i)$, and then $\sum_{i\in\mathcal{B}} \frac{\rho_i}{1-\rho_i}$ is the total number of flows in the system [KdVYV12]. According to the Little's law, minimising the average number of flows is equivalent to minimising the average delay experienced by a typical traffic flow.

5.2.2.2 Cost Function of On-Grid Energy Consumption

The green traffic load is defined as the maximum traffic load that can be supported by the green energy harvested from renewable energy sources. With G_i denoted as the amount of green energy available to BS i, based on equation (5.5), the green traffic load

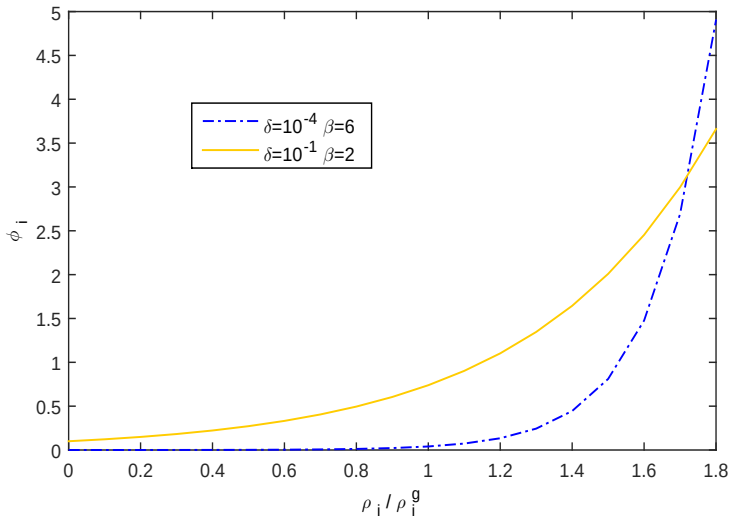


Figure 5.3: Cost of on-grid energy consumption with different values of β and δ.

of BS i is derived as

$$\rho_i^g = \max\left(\varepsilon, \min\left(\frac{G_i - E_i^s}{\Delta_i p_i \tau}, 1-\varepsilon\right)\right), \tag{5.10}$$

note that ε is an arbitrarily small positive constant to ensure $0 < \rho_i^g < 1$.

According to (5.6), on-grid energy is only consumed when green energy is not sufficient. Thus when the traffic load of BS i exceeds the amount of green traffic load, that is $\rho_i > \rho_i^g$, on-grid energy will be consumed, which leads to the increase in the cost of on-grid energy consumption. Otherwise, the cost of on-grid energy consumption will stay trivial. In doing so, the cost function of on-grid energy consumption is designed as

$$\phi(\boldsymbol{\rho}) = \sum\nolimits_{i\in\mathcal{B}} \phi_i(\rho_i) = \sum\nolimits_{i\in\mathcal{B}} \delta \exp\left(\beta \frac{\rho_i}{\rho_i^g}\right), \tag{5.11}$$

where β represents the network sensitivity towards on-grid energy consumption ($\beta > 0$), and δ aims to adjust the value of cost function ($\delta > 0$). Fig. 5.3 shows curves of cost function of on-grid energy consumption versus ρ_i/ρ_i^g with different values of β and δ. It is shown that with proper values of β and δ, such as $\beta = 6$ and $\delta = 10^{-4}$, which is adopted in this section, the cost function of on-grid energy consumption has the following property:

when $\rho_i/\rho_i^g > 1$, $\phi_i(\rho_i)$ increases exponentially with the rise of ρ_i; when $0 < \rho_i/\rho_i^g < 1$, $\phi_i(\rho_i)$ remains almost zero. Such a property encourages the good use of the green energy harvested from renewable energy sources, in order to reduce on-grid energy consumption.

5.2.3 User Association Algorithm for Delay-Energy Tradeoffs

In this section, the user association algorithm is proposed to achieve the global optimum in minimising the total system cost $f(\boldsymbol{\rho})$. The proposed user association algorithm is implemented in an iterative manner: BSs periodically measure and advertise their loads, and then users make user association decision based on the advertised information to minimise $f(\boldsymbol{\rho})$. The BS and user sides update iteratively until convergence. The proposed user association algorithm is totally distributed, and does not require any centralised computation. As such, it will not incur algorithmic complexity issue here.

In order to guarantee convergence, it is assumed that spatial load distributions in HetNets area are temporally stationary, and the time scale of users making user association decision is faster than that of BSs advertising their loads. In this case, BSs advertise their load conditions after the system remains stationary. It is also assumed that the available green energy of every BS is constant during the period of determining user association.

The proposed user association algorithm consists of two parts.

User side: At the beginning of k-th time slot, users get the traffic loads $\boldsymbol{\rho}^{(k)}$ of all BSs via the broadcast. And then the user at location x chooses the optimal BS by

$$j^{(k)}(x) = \arg\max_{i \in \mathcal{B}} \frac{r_i(x)}{\left(1-\rho_i{}^{(k)}\right)^{-2} + \frac{\omega\beta\delta}{\rho_i^g} \exp\left(\beta \frac{\rho_i{}^{(k)}}{\rho_i^g}\right)}. \tag{5.12}$$

Then the user association indicator is updated by

$$y_i{}^{(k)}(x) = \begin{cases} 1, & \text{if } i = j^{(k)}(x) \\ 0, & \text{otherwise.} \end{cases} \tag{5.13}$$

And $y_i^{(k)}(x)$ will be broadcasted to all BSs.

BS side: The updated user association indicators from the user side will change loads of BSs, and thus during the k-th period, the new traffic load of BS i is given by

$$T_i\left(\rho_i{}^{(k)}\right) = \min\left(\int_{\mathcal{L}} \frac{\lambda(x)\,\mu(x)\,y_i{}^{(k)}(x)}{r_i(x)} dx,\ 1-\varepsilon\right). \tag{5.14}$$

Based on the derived $T_i\left(\rho_i{}^{(k)}\right)$, BS i updates the next advertising traffic load as [KdVYV12]

$$\rho_i{}^{(k+1)} = \theta\rho_i{}^{(k)} + (1-\theta)\,T_i\left(\rho_i{}^{(k)}\right), \quad \forall i \in \mathcal{B}, \tag{5.15}$$

where $0 \leq \theta < 1$ is an exponential averaging parameter.

The following provides proof on optimality and convergence of the proposed user association algorithm.

Lemma 1: A unique optimal $\boldsymbol{\rho}*$ exists to minimise $f(\boldsymbol{\rho}) = \phi(\boldsymbol{\rho}) + \omega\varphi(\boldsymbol{\rho})$, when $\boldsymbol{\rho}$ is defined on $\tilde{\mathcal{F}}$.

Proof. The objective function $f(\boldsymbol{\rho}) = \phi(\boldsymbol{\rho}) + \omega\varphi(\boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$ when $\boldsymbol{\rho}$ is defined on $\tilde{\mathcal{F}}$, since $\nabla^2 f(\boldsymbol{\rho}) > 0$ when $\boldsymbol{\rho} \in \tilde{\mathcal{F}}$. As such, there exists a unique optimal $\boldsymbol{\rho}*$ that minimises $f(\boldsymbol{\rho})$. □

We denote $\boldsymbol{\rho}^{(k)} = \left(\rho_1{}^{(k)}, \cdots, \rho_{|\mathcal{B}|}{}^{(k)}\right)$ and $\mathbf{T}\left(\boldsymbol{\rho}^{(k)}\right) = \left(T_1\left(\rho_1{}^{(k)}\right), \cdots, T_{|\mathcal{B}|}\left(\rho_{|\mathcal{B}|}{}^{(k)}\right)\right)$.

Lemma 2: When $\boldsymbol{\rho}^{(k)} \neq \boldsymbol{\rho}*$, $\mathbf{T}\left(\boldsymbol{\rho}^{(k)}\right) - \boldsymbol{\rho}^{(k)}$ is a descent direction of $f\left(\boldsymbol{\rho}^{(k)}\right)$.

Proof. This lemma can be proved by deriving $\left\langle \nabla f\left(\boldsymbol{\rho}^{(k)}\right), \mathbf{T}\left(\boldsymbol{\rho}^{(k)}\right) - \boldsymbol{\rho}^{(k)} \right\rangle \leq 0$. Let

$y_i(x)$ and ${y_i}^T(x)$ be the user association indicators of BS i that result in the traffic load ${\rho_i}^{(k)}$ and $T_i\left({\rho_i}^{(k)}\right)$, respectively.

$$\begin{aligned}
&\left\langle \nabla f\left(\boldsymbol{\rho}^{(k)}\right), \mathbf{T}\left(\boldsymbol{\rho}^{(k)}\right) - \boldsymbol{\rho}^{(k)} \right\rangle \\
&= \sum_{i\in\mathcal{B}} \left(T_i\left({\rho_i}^{(k)}\right) - {\rho_i}^{(k)}\right)\left(\left(1-{\rho_i}^{(k)}\right)^{-2} + \frac{\omega\beta\delta\exp\left(\beta\frac{{\rho_i}^{(k)}}{\rho_i^g}\right)}{\rho_i^g}\right) \\
&= \sum_{i\in\mathcal{B}} \Big[\int_{\mathcal{L}} \lambda(x)\,\mu(x)\,(r_i(x))^{-1}\left({y_i}^T(x) - y_i(x)\right)dx \\
&\times \left(\left(1-{\rho_i}^{(k)}\right)^{-2} + \omega\beta\delta(\rho_i^g)^{-1}\exp\left(\beta\frac{{\rho_i}^{(k)}}{\rho_i^g}\right)\right)\Big] \\
&= \int_{\mathcal{L}} \lambda(x)\,\mu(x) \sum_{i\in\mathcal{B}} \Big[(r_i(x))^{-1}\left({y_i}^T(x) - y_i(x)\right) \\
&\times \left(\left(1-{\rho_i}^{(k)}\right)^{-2} + \omega\beta\delta(\rho_i^g)^{-1}\exp\left(\beta\frac{{\rho_i}^{(k)}}{\rho_i^g}\right)\right)\Big]\,dx.
\end{aligned} \tag{5.16}$$

Note that

$$\begin{aligned}
&\sum_{i\in\mathcal{B}} \frac{{y_i}^T(x)\left(\left(1-{\rho_i}^{(k)}\right)^{-2} + \omega\beta\delta\left(\rho_i^g\right)^{-1}\exp\left(\beta\frac{{\rho_i}^{(k)}}{\rho_i^g}\right)\right)}{r_i(x)} \\
&\leq \sum_{i\in\mathcal{B}} \frac{y_i(x)\left(\left(1-{\rho_i}^{(k)}\right)^{-2} + \omega\beta\delta\left(\rho_i^g\right)^{-1}\exp\left(\beta\frac{{\rho_i}^{(k)}}{\rho_i^g}\right)\right)}{r_i(x)}
\end{aligned} \tag{5.17}$$

holds, since the user association indicator ${y_i}^T(x)$ derived from equation (5.12) (5.13) maximises the value of $r_i(x)\left(\left(1-{\rho_i}^{(k)}\right)^{-2} + \omega\beta\delta(\rho_i^g)^{-1}\exp\left(\beta\frac{{\rho_i}^{(k)}}{\rho_i^g}\right)\right)^{-1}$, for all $i \in \mathcal{B}$. Hence $\left\langle \nabla f\left(\boldsymbol{\rho}^{(k)}\right), \mathbf{T}\left(\boldsymbol{\rho}^{(k)}\right) - \boldsymbol{\rho}^{(k)} \right\rangle \leq 0$. □

Theorem 1 (Convergence): The traffic load $\boldsymbol{\rho}$ converges to $\boldsymbol{\rho}* \in \mathcal{F}$.

Proof. Since ${\rho_i}^{(k+1)} - {\rho_i}^{(k)} = \theta{\rho_i}^{(k)} + (1-\theta)\,T_i\left({\rho_i}^{(k)}\right) - {\rho_i}^{(k)} = (1-\theta)\left(T_i\left({\rho_i}^{(k)}\right) - {\rho_i}^{(k)}\right)$, and $0 \leq \theta < 1$, $\boldsymbol{\rho}^{(k+1)} - \boldsymbol{\rho}^{(k)}$ is also a descent direction of $f\left(\boldsymbol{\rho}^{(k)}\right)$ according to **Lemma 2**. Then based on **Lemma 1** where $f\left(\boldsymbol{\rho}^{(k)}\right)$ is a convex function and is lower bounded by 0, it comes to the conclusion that $f\left(\boldsymbol{\rho}^{(k)}\right)$ converges. Suppose $f\left(\boldsymbol{\rho}^{(k)}\right)$ converges to some point other than $f(\boldsymbol{\rho}*)$, then $\boldsymbol{\rho}^{(k+1)}$ produces a descent direction again, which means $f\left(\boldsymbol{\rho}^{(k)}\right)$ can further decrease in the next iteration. This contradicts the convergence assumption, and thus $\boldsymbol{\rho}^{(k)}$ converges to $\boldsymbol{\rho}*$. Since $\boldsymbol{\rho}^{(k)}$ is derived based on (5.12), (5.13), (5.14), (5.15), where ${y_i}^{(k)}(x) \in \{0,1\}$, $\boldsymbol{\rho}*$ is in the feasible set $\mathcal{F}$. □

Remark 2: The computational complexity of the user association algorithm for an individual user is $\mathcal{O}(|\mathcal{B}|)$ for each iteration. According to [KdVYV12], although the convergence speed depends on the value of θ, fixed θ close to 1 generally works well for the convergence. However, how to optimise θ is beyond my research scope. In the simulation, the value of θ is set as 0.98, and simulation results have shown the proposed user association algorithm converges quickly to the optimum, within 60 iterations.

Theorem 2: Suppose the feasible set $\mathcal{F}$ is not empty, and the traffic load $\boldsymbol{\rho}$ converges to $\boldsymbol{\rho}*$, the user association corresponding to $\boldsymbol{\rho}*$ minimises $f(\boldsymbol{\rho})$, which is the optimal solution of the user association problem (5.8).

Proof. Let $\boldsymbol{y}*$ and $\boldsymbol{y}$ be the user association indicators corresponding to $\boldsymbol{\rho}*$ and $\boldsymbol{\rho}$, respectively. Based on **Theorem 1**, $\boldsymbol{\rho}*$ is in the feasible set $\mathcal{F}$ with the corresponding deterministic user association $\boldsymbol{y}* = \{y_i*(x) \,|y_i*(x) \in \{0,1\}, \forall i, \forall x\}$.

Since $f(\boldsymbol{\rho})$ is a convex function over $\boldsymbol{\rho}$, proving the theorem is equivalent to prove

$$\langle \nabla f(\boldsymbol{\rho}*)\,|_{\boldsymbol{\rho}=\boldsymbol{\rho}*}, \boldsymbol{\rho} - \boldsymbol{\rho}* \rangle \geq 0. \tag{5.18}$$

$$\begin{aligned}
&\langle \nabla f(\boldsymbol{\rho}*)\,|_{\boldsymbol{\rho}=\boldsymbol{\rho}*}, \boldsymbol{\rho} - \boldsymbol{\rho}* \rangle \\
&= \sum_{i\in\mathcal{B}} \left((1-\rho_i*)^{-2} + \frac{\omega\beta\delta \exp\left(\beta\frac{\rho_i*}{\rho_i^g}\right)}{\rho_i^g}\right)(\rho_i - \rho_i*) \\
&= \sum_{i\in\mathcal{B}} \left((1-\rho_i*)^{-2} + \frac{\omega\beta\delta \exp\left(\beta\frac{\rho_i*}{\rho_i^g}\right)}{\rho_i^g}\right) \int_{\mathcal{L}} \frac{\lambda(x)\,\mu(x)\,(y_i(x) - y_i*(x))}{r_i(x)} dx \\
&= \int_{\mathcal{L}} \lambda(x)\,\mu(x) \sum_{i\in\mathcal{B}} \left[(r_i(x))^{-1}\,(y_i(x) - y_i*(x))\,\left((1-\rho_i*)^{-2} + \omega\beta\delta(\rho_i^g)^{-1}\exp\left(\beta\frac{\rho_i*}{\rho_i^g}\right)\right)\right] dx.
\end{aligned} \tag{5.19}$$

Since the optimal user association indicator is determined by

$$y_i*(x) = \begin{cases} 1, \text{ if } i = \arg\max\limits_{i\in\mathcal{B}} \frac{r_i(x)}{(1-\rho_i*)^{-2}+\frac{\omega\beta\delta}{\rho_i^g}\exp\left(\beta\frac{\rho_i*}{\rho_i^g}\right)} \\ 0, \text{ otherwise} \end{cases}, \tag{5.20}$$

we have

$$\sum_{i\in\mathcal{B}}\left[\left(r_i\left(x\right)\right)^{-1}\left(y_i\left(x\right)-y_i*\left(x\right)\right)\left(\left(1-\rho_i*\right)^{-2}+\omega\beta\delta(\rho_i^g)^{-1}\exp\left(\beta\frac{\rho_i*}{\rho_i^g}\right)\right)\right]\geq 0. \tag{5.21}$$

Hence $\langle\nabla f\left(\boldsymbol{\rho}*\right)|_{\boldsymbol{\rho}=\boldsymbol{\rho}*}, \boldsymbol{\rho}-\boldsymbol{\rho}*\rangle \geq 0$. □

5.2.4 Admission Control

Up till now, the problem formulation in Section 5.2.2 and the proposed algorithm in Section 5.2.3 assume the condition where the user association problem (5.8) is feasible, that is $\rho_i \leq 1-\varepsilon, \forall i \in \mathcal{B}$. However when the traffic loads are high, the user association problem (5.8) will not be feasible, and thus the admission control is required. In this section, the admission control is investigated with the objective to minimise the system cost which includes the cost of blocking traffic. Here, the blocking cost is assumed to be proportional to the amount of blocked traffic. As the traffic blocking affects users' satisfaction, such admission control policy is able to reflect the business concern of network operators.

It is assumed that the blocked traffic is routed to the *null* BS. $\mathcal{B}_0$ is used to denote all BSs including the *null* BS, and $\boldsymbol{\rho}_0 = \left(\rho_0, \rho_1, \ldots, \rho_{|\mathcal{B}|}\right)$. Note that ρ_0 is defined as $\rho_0 = \int_{\mathcal{L}} \lambda\left(x\right)\mu\left(x\right)\left(1-\sum_{i\in\mathcal{B}} y_i\left(x\right)\right)dx$, where $1-\sum_{i\in\mathcal{B}} y_i\left(x\right)$ means if a user does not associate with any BS, this user is blocked. ρ_0 is the total amount of traffic that is

blocked and ρ_0 can be larger than 1. Then the feasible set $\tilde{\mathcal{F}}_0$ including ρ_0 is defined as

$$\begin{aligned}\tilde{\mathcal{F}}_0 = \{\boldsymbol{\rho_0}|\ \rho_0 = \textstyle\int_{\mathcal{L}} \lambda(x)\,\mu(x)\left(1-\sum_{i\in\mathcal{B}} y_i(x)\right)dx,\\ \rho_i = \textstyle\int_{\mathcal{L}} \varrho_i dx, 0 \le \rho_i \le 1-\varepsilon, \forall i \in \mathcal{B},\\ \textstyle\sum_{i\in\mathcal{B}_0} y_i(x) = 1, 0 \le y_i(x) \le 1, \forall x, \forall i \in \mathcal{B}_0\}.\end{aligned} \tag{5.22}$$

The optimisation problem is given by

$$\min_{\boldsymbol{\rho}_0}\left\{f(\boldsymbol{\rho}_0) = \varphi(\boldsymbol{\rho}) + \omega\phi(\boldsymbol{\rho}) + \alpha\rho_0 | \boldsymbol{\rho}_0 \in \tilde{\mathcal{F}}_0\right\}, \tag{5.23}$$

where $\alpha\rho_0$ is the cost of blocking traffic and α reflects the blocking cost per bit. With the similar proof in the Section 5.2.2, and Section 5.2.3, it is easy to conclude that the user association problem (5.23) is also a convex optimisation problem defined on a convex set. Hence the user association algorithm is similar as the algorithm in Section 5.2.3, only with the revised equation (5.12),

$$j^{(k)}(x) = \arg\max_{i\in\mathcal{B}_0} \upsilon_i(x), \tag{5.24}$$

where

$$\upsilon_i(x) = \begin{cases} \alpha^{-1}, & \text{if } i = 0 \\ \frac{r_i(x)}{\left(1-\rho_i{}^{(k)}\right)^{-2} + \frac{\omega\beta\delta}{\rho_i^g}\exp\left(\beta\frac{\rho_i{}^{(k)}}{\rho_i^g}\right)}, & \text{if } i \in \{1,2,\cdots,|\mathcal{B}|\}, \end{cases} \tag{5.25}$$

here α^{-1} acts as the threshold to determine whether a particular user is blocked or not. More specifically, a BS blocks users that do not have good performance compared with the threshold α^{-1}. It is also observed that the threshold α^{-1} is the inverse of the blocking cost per bit α. Therefore, the user is less likely to be blocked with higher blocking cost, and vice versa.

Table 5-B: Green Energy Distribution in Different Scenarios

Scenario	**MBS**	**PBS1**	**PBS2**	**PBS3**
Scenario 1	0.870kWh	0.033kWh	0.032kWh	0.033kWh
Scenario 2	0.837kWh	0.033kWh	0.032kWh	0.033kWh

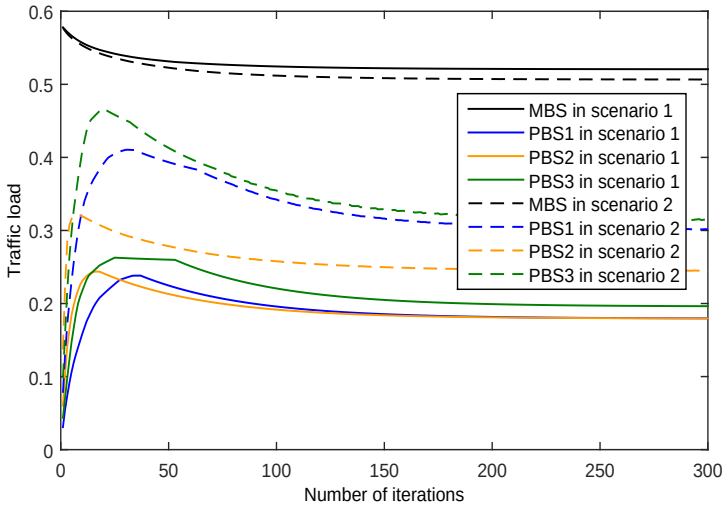


Figure 5.4: Traffic load in different scenarios with different distributions of green energy.

5.2.5 Simulation Results and Conclusions

5.2.5.1 Simulation Results

The performance of the proposed user association (proposed UA) is evaluated by the simulation platform presented in Section 5.1.2.

First the weight ω between cost of average traffic delay and on-grid energy consumption is set as 10^{-1}, and traffic arrival rate as 1. Fig. 5.4 shows curves of traffic loads in the proposed UA in different scenarios with different distributions of green energy harvested from renewable energy sources. Table 5-B specifies the green energy distribution in different scenarios. Fig. 5.4 shows that compared with scenario 2, the traffic load of MBS is higher in scenario 1, where MBS has larger green energy. This figure indicates that the proposed UA is able to adapt traffic loads among BSs along with distributions of green energy, in order to make good use of renewable energy and reduce on-grid energy

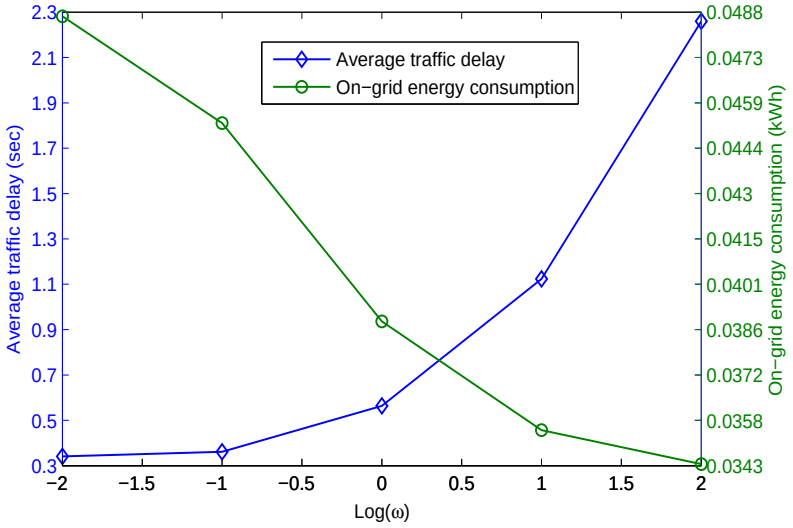


Figure 5.5: Average traffic delay and on-grid energy consumption versus different values of weight ω.

consumption. Fig. 5.4 also reveals the convergence of traffic loads in the proposed UA. It is shown that the proposed UA converges quickly to the global optimum.

Then the performance of the proposed UA is further evaluated in in scenario 1. Fig. 5.5 demonstrates the average traffic delay and on-grid energy consumption versus different values of weight ω, where with the increase of weight ω, the average traffic delay rises and the on-grid energy consumption decreases. The larger weight ω will lead to more emphasis on the energy saving in order to minimise on-grid energy consumption, while the smaller weight ω will attach more importance to the average traffic delay minimisation. This figure indicates the weight ω in the proposed UA is able to adjust the tradeoff between average traffic delay and on-grid energy consumption.

Additionally, the performance of the proposed UA is compared with the delay-optimal user association algorithm (delay-optimal UA) in [KdVYV12]. In the delay-optimal UA, user association is determined to minimise the average traffic delay. Fig. 5.6 and Fig. 5.7 show the average traffic delay and on-grid energy consumption versus different values of traffic arrival rate λ, respectively. Fig. 5.6 indicates the proposed UA is slightly inferior to the delay-optimal UA in terms of average traffic delay. However the proposed UA

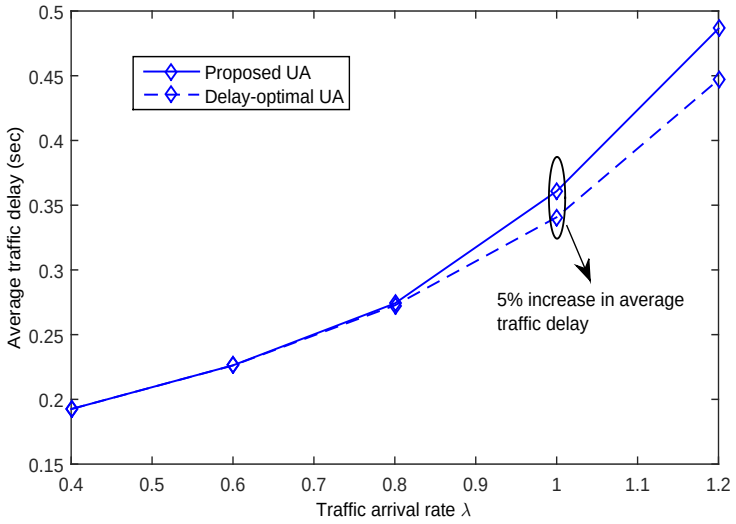


Figure 5.6: Average traffic delay versus different values of traffic arrival rate λ.

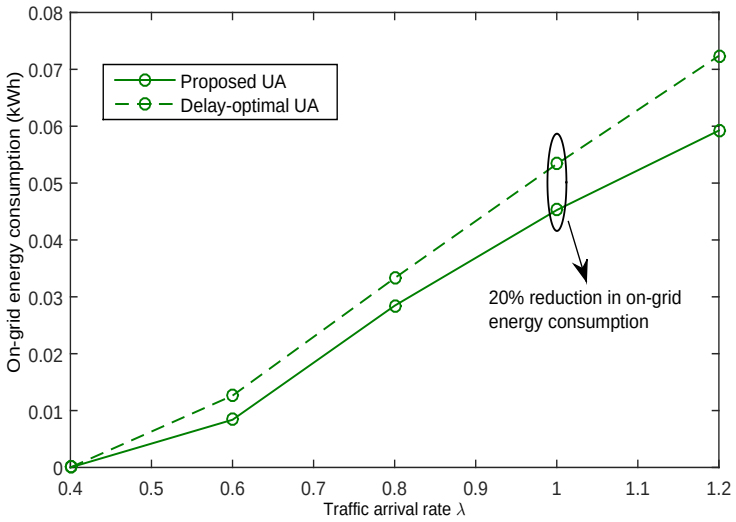


Figure 5.7: On-grid energy consumption versus different values of traffic arrival rate λ.

obtains significant improvement in on-grid energy saving shown in Fig. 5.7. Hence, the proposed UA achieves the comparable average traffic delay compared with the delay-optimal UA, while substantially reducing the on-grid energy consumption.

Finally, the performance of the proposed UA with admission control is evaluated in the heavy traffic load condition with traffic arrival rate as 1.4. Fig. 5.8 validates that the admission control enables the proposed UA to work in heavy traffic load condition. It also

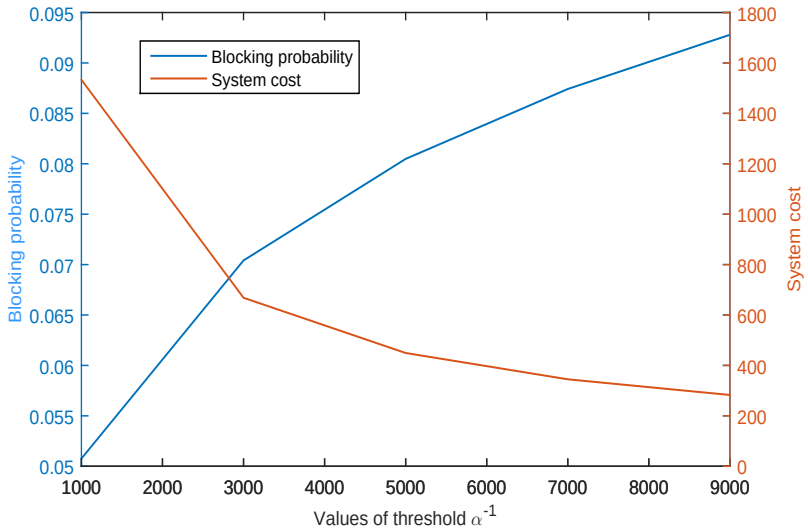


Figure 5.8: System cost and blocking probability with different values of threshold α^{-1}.

indicates that the rise of the threshold degrades the network performance by increasing the blocking probability, as well as reduces the system cost which is the sum of cost of average traffic delay, cost of on-grid energy consumption and cost of blocking traffic. Hence, in practical system, the threshold α^{-1} should be carefully chosen considering the tradeoff between blocking probability and system cost.

5.2.5.2 Conclusions

In this section, an optimal user association algorithm was proposed for delay and energy consumption tradeoffs in HetNets with hybrid energy sources, where all BSs are assumed to be powered by both power grid and renewable energy sources. With the convex optimisation problem formulation, it was proven that the proposed user association algorithm converged to the global optimum which minimised the weighted sum of cost of average traffic delay and cost of on-grid energy consumption. The proposed user association algorithm allows for a flexible tradeoff between average traffic delay and on-grid energy consumption by adjusting the value of weight ω. Admission control was also addressed to ensure the proposed user association works in the heavy traffic load condition. Simulation results validate the merits of the proposed user association algorithm not only in

adapting loads of BSs along with distributions of green energy, but also in substantially reducing on-grid energy consumption and achieving comparable average traffic delay compared to the existing algorithm which aims to minimise the average traffic delay.

5.3 Two-Dimensional Optimisation on User Association and Green Energy Allocation for HetNets with Hybrid Energy Sources

Section 5.3 considers the two-dimensional optimisation on user association and green energy allocation to optimise the on-grid energy consumption in HetNets with hybrid energy sources. In this section, it is assumed that each BS is equipped with a rechargeable battery with maximum capacity as B_i^{max} to store the harvested and residual green energy. According to [GWZ13], rechargeable battery can be modelled by an ideal linear model, where changes in the energy stored are linearly related to amounts of energy harvest or spent, provided that the maximum battery capacity is not exceeded. Note that small BSs, such as PBSs, may have smaller energy harvesting rates and battery capabilities than those of MBSs.

5.3.1 Motivation

In the hybrid energy powered scenario, on-grid energy is consumed only when the total energy consumption exceeds the allocated green energy. As such, the reduction of on-grid energy consumption relies on the optimisation of energy consumption and green energy allocation. However, the renewable energy always exhibits temporal dynamic, such as the daily solar energy generation in a given area peaks around noon, and bottoms during the night. The energy consumption of BSs depends on the mobile traffic load, which also shows both temporal and spatial dynamics [PLL$^+$11]. Thus the optimisation of on-grid energy consumption involves the optimisation in both time and space dimensions.

Furthermore, in contrast with the existing research which considers the total on-grid energy consumption minimisation only, the two-dimensional optimisation in this section aims to minimise both total and peak on-grid energy consumptions. This can be justified by the fact that in real networks, the high peak on-grid energy consumption will translate into high OPEX for mobile network operators, as electric systems need to maintain sufficient capacity of on-grid energy generation to meet the expected peak on-grid energy consumption plus a reserve margin [Dep14]. In addition, extra costs of high-capacity equipments are usually covered by industrial consumers, such as mobile network operators, thereby resulting in the increasing OPEX for mobile network operators. In this sense, it is not sufficient to only reduce the total on-grid energy consumption. The minimisation of peak on-grid energy consumption is also imperative.

The main contributions of the two-dimensional optimisation are presented as follows.

- Distinct with the existing research which focuses on the minimisation of total on-grid energy consumption only, the two-dimensional optimisation problem is formulated to lexicographically minimise the on-grid energy consumption in HetNets with hybrid energy sources, thereby reducing both total and peak on-grid energy consumptions.

- Taking advantage of the time scale separation assumption, the problem is decomposed into two sub-optimisation problems without loss of optimality of the original optimisation problem. The user association optimisation in space dimension is first formulated via convex optimisation to minimise total energy consumption through distributing the traffic appropriately across different BSs in a certain time slot. Then the green energy allocation is optimised across different time slots for an individual BS to lexicographically minimise the on-grid energy consumption, based on the traffic load determined by the first sub-optimisation problem.

- To solve the two-dimensional optimisation problem, a low complexity optimal offline algorithm with infinite battery capacity is proposed by assuming the non-

causal green energy and traffic knowledge, that amounts to the harvested green energy and traffic of all time slots known as priori. The optimality of the proposed optimal offline algorithm is theoretically proven, and the effect of the proposed algorithm is demonstrated by simulation results.

- In practice, the amounts of harvested green energy and traffic are random in nature and cannot be predicted precisely in advance. In this case, online algorithms relying on the information of previous and current time slots are required. The proposed optimal offline algorithm provides useful performance upper bound for the more practical online algorithms, and sheds light on the design of online algorithms. Thus inspired by the proposed optimal offline algorithm, some heuristic online algorithms with finite battery capacity are developed by utilising causal green energy and traffic information, and their performance is evaluated via simulations.

5.3.2 Problem Formulation

Due to disadvantages of high peak on-grid energy consumption, the objective is to not only reduce the total on-grid energy consumption but also achieve the time-fair on-grid energy consumption, that is, to make the on-grid energy consumption evenly distributed with respect to time as much as possible. This can be achieved by using lexicographic minimisation which is defined as below.

Definition 1. *Let* $\mathbf{A_1}$ *and* $\mathbf{A_2}$ *be two resource allocations. A resource vector* $\mathbf{L^A}$ *is a sorted resource vector of a resource allocation, if* $\mathbf{L^A}$ *is the result of sorting* $\mathbf{A}$ *in non-increasing order, and* L_i^A *is the i-th element in* $\mathbf{L^A}$*. We say* $\mathbf{A_1}$=$\mathbf{A_2}$ *if* $\mathbf{L^{A_1}}$=$\mathbf{L^{A_2}}$*,* $\mathbf{A_1}$<$\mathbf{A_2}$ *if there exists an* i *such that* $L_i^{A_1} < L_i^{A_2}$ *and* $L_j^{A_1} = L_j^{A_2}, \forall j < i$*, and* $\mathbf{A_1}$>$\mathbf{A_2}$ *otherwise.* $\mathbf{A}^*$ *is the optimal lexicographic minimisation resource allocation if there is no other resource allocation* $\mathbf{A}'$<$\mathbf{A}^*$*.*

Note that lexicographic optimisation is a well-established method for radio resource allocation control in wireless networks, see [RLB07, RT10] and references therein.

We denote $\boldsymbol{G} = \{G_i(t) | \forall i, \forall t\}$ and $\boldsymbol{\rho} = \{\rho_i(t) | \forall i, \forall t\}$. Thus, in this section, the problem is to find optimal load of BSs $\boldsymbol{\rho}$ and optimal green energy allocation $\boldsymbol{G}$ in order to lexicographically minimise the on-grid energy consumption, which is given by

$$\mathbf{P1}: \quad \underset{\boldsymbol{\rho},\mathrm{G}}{\text{Lexicographically minimise}} \left\{ \sum_{i\in\mathcal{B}} E_i^{grid}(1), \cdots, \sum_{i\in\mathcal{B}} E_i^{grid}(t), \cdots, \sum_{i\in\mathcal{B}} E_i^{grid}(|\mathcal{T}|) \right\}, \tag{5.26}$$

$$\text{s.t.}: R_i(1) = B_i^0, \ \forall i \tag{5.27}$$

$$R_i(t) = R_i(t-1) + Q_i(t-1) - G_i(t-1), \ \forall i, t \geq 2 \tag{5.28}$$

$$G_i(t) \leq R_i(t) + Q_i(t), \ \forall i, \ \forall t \tag{5.29}$$

$$R_i(t) + Q_i(t) \leq B_i^{\max}, \ \forall i, \ \forall t \tag{5.30}$$

$$\boldsymbol{\rho} \in \mathcal{F}, \ \forall t, \tag{5.31}$$

where $Q_i(t)$ is the amount of green energy harvested by BS i at t-th time slot. $R_i(t)$ is the residual energy left from previous time slots, and it is available at t-th time slot. $B_i^{\max}$ is the maximum battery capacity of BS i. $G_i(t)$ is the green energy allocation of BS i at t-th time slot. It is assumed that the harvested green energy $Q_i(t)$ arrives at the beginning of t-th time slot, and the initial energy stored in the battery is B_i^0. (5.28) represents the 'storage evolution' dynamics. (5.29) is the energy causality constraint, that is BS cannot consume more green energy than it has stored. (5.30) means excessive energy cannot be stored in the rechargeable battery.

The lexicographic minimisation of on-grid energy consumption **P1** aims to find the optimal green energy allocation $\boldsymbol{G}$, as well as determine which BS each user should associate with, or equivalently, to find the optimal loads of BSs $\boldsymbol{\rho}$, in order to minimise both total and peak on-grid energy consumptions. The on-grid energy is consumed only when the total energy consumption exceeds the green energy allocated in each time slot. As such, the reduction of on-grid energy consumption depends on the optimisation of both the energy consumption and green energy allocation. According to equation (5.5), the energy consumption of BS is decided by the traffic load of BS, which exhibits temporal and spacial dynamics. In this sense, in order to reduce the total energy consumption

throughout the whole network, user association decision should be made to distribute the traffic load appropriately among all BSs, thereby minimising the energy consumption. The renewable energy also shows temporal dynamic, such as the daily solar energy generation in a given area peaks around noon, and bottoms during the night. As such, to minimise both total and peak on-grid energy consumptions, the green energy allocation should be optimised among all time slots $\mathcal{T}$. In other words, the lexicographic minimisation of on-grid energy consumption **P1** involves optimisation in both time and space dimensions. In the space dimension, traffic load among BSs should be distributed appropriately within the whole network. In the time dimension, the green energy allocation across different time slots should be optimised.

Hence, solving **P1** is very challenging as the highly complex coupling of user association and green energy allocation. For analysis tractability, an assumption on the time-scale separation is made that traffic request arrival and departure process and the corresponding user association process are much faster than the period on which the green energy allocation across different time slots is determined. From statistics in real networks [PLL$^+$11, NRE], the traffic pattern, e.g., traffic distribution, and green energy generation rate vary over time, but could be assumed almost constant during a certain period, e.g., one hour. Since the time scale for determining green energy allocation is similar to the order of traffic pattern and green energy generation rate changing, it is definitely much larger than that of traffic request arrival and departure process, e.g., typically less than several minutes [SKYK11].

With this in mind, it is assumed that both the traffic pattern and green energy generation rate are stationary during one time slot. Furthermore, shown in (5.6), $E_i^{grid}(t)$ reduces with decreasing $E_i(t)$, where $E_i(t)$ is determined by $\rho_i(t)$, the traffic load of BS i at t-th time slot. Furthermore, the value of $E_i(t)$ is independent with the value of green energy allocation $G_i(t)$. Therefore the optimisation problem is decomposed into two sub-optimisation problems. In space dimension, the traffic load is distributed appropriately across different BSs $\mathcal{B}$ in a certain time slot to minimise the total energy

consumption $\sum_{i\in\mathcal{B}} E_i(t)$. While in time dimension, the green energy allocation $\boldsymbol{G}$ is optimised across different time slots $\mathcal{T}$ in order to lexicographically minimise the on-grid energy consumption, based on the traffic load derived from the optimisation in space dimension. Note that solutions for these two decomposed sub-optimisation problems are also optimal for the original problem **P1**, which will be made clear in the proof of **Theorem 4** in Section 5.3.3.3.

5.3.2.1 Space Dimension: User Association Optimisation

The user association optimisation in space dimension is to determine which BS each user should associate with, or equivalently, to find the optimal traffic load $\boldsymbol{\rho}$ for any given time slot. It aims to minimise the total energy consumption of all BSs during one time slot, while guaranteeing QoS requirements for all users in a sense that all users are served with the required traffic amount, which is given by

$$\begin{aligned} &\min_{\boldsymbol{\rho}} \quad \textstyle\sum_{i\in\mathcal{B}} \left[E_i(t)\right] \\ &\text{s.t. : (5.31)} \end{aligned} . \tag{5.32}$$

Since $y_i(x) \in \{0,1\}$, $\mathcal{F}$ in (5.4) is not a convex set. Similar to the approach used in Section 5.2.2, in order to facilitate the convex optimisation problem formulation, $y_i(x) \in \{0,1\}$ is relaxed to $0 \le y_i(x) \le 1$, where $y_i(x)$ specifies the probability that the user at location x is associated with BS i. The time slot index t is omitted here for simplicity, and then the updated set $\tilde{\mathcal{F}}$ of feasible loads of BSs is

$$\begin{aligned} \tilde{\mathcal{F}} = \{\boldsymbol{\rho} |\ & \rho_i = \textstyle\int_{\mathcal{L}} \varrho_i dx, \\ & 0 \le \rho_i \le 1-\varepsilon, \textstyle\sum_{i\in\mathcal{B}} y_i(x) = 1, \\ & 0 \le y_i(x) \le 1, \forall x, \forall i\}. \end{aligned} \tag{5.33}$$

The authors in [KdVYV12] have proved that the set $\tilde{\mathcal{F}}$ is convex.

Then a penalty function $L_i(\rho_i(t))$ is introduced to the original sub-optimisation problem (5.32), and the user association optimisation in space dimension is formulated as a convex optimisation given by

$$\begin{aligned} \mathbf{P1.1}: \quad & \min_{\boldsymbol{\rho}} \quad \textstyle\sum_{i\in\mathcal{B}} \left[E_i(t) + L_i(\rho_i(t))\right] \\ & \text{s.t.}: \boldsymbol{\rho} \in \tilde{\mathcal{F}}, \ \forall t, \end{aligned} \tag{5.34}$$

By adding penalty $L_i(\rho_i(t))$ into the objective, the traffic could be balanced among BSs, which avoids cells getting too congested. Furthermore, such load balancing benefits the QoS provision by reducing the average traffic delay. Although there may be other methods to design the penalty function, in this section, the penalty is defined as following

$$L_i(\rho_i(t)) = \omega_i \log\left(\frac{1}{1-\rho_i(t)}\right), \tag{5.35}$$

where ω_i is the weight to adjust the significance of the penalty of BS i. Larger value of traffic load $\rho_i(t)$ will lead to higher penalty value. Due to the low transmit power and limited capacity of picocells, picocells have higher chance to become the early capacity bottleneck. Thus the weight of picocell always has higher value than that of macrocell. In general, the value of weight can be selected as a composite tradeoff among signal quality, spectrum efficiency and load balancing needs, and it can also be adapted in real time to address dynamic changes in the network. The effect of different values of ω_i will be demonstrated in the Section 5.3.5. It is worthwhile mentioning that as the value of ω_i goes to zero, the objective function of the modified sub-optimisation problem **P 1.1** given in (5.34) is asymptotically equivalent to the objective function of the original sub-optimisation problem without penalty function given in (5.32).

5.3.2.2 Time Dimension: Green Energy Allocation Optimisation

Based on the traffic load derived from the user association optimisation in space dimension, the green energy allocation in time dimension is to optimise the green energy

allocation across different time slots to lexicographically minimise the on-grid energy consumption. As there is no energy transfer among BSs in this model, the lexicographical minimisation of on-grid energy consumption of the whole network can be achieved by lexicographically minimising the on-grid energy consumption of every individual BS separately, which is formulated as

$$\begin{aligned} &\mathbf{P1.2}: \quad \text{Lexicographically minimise} \\ &\underset{\mathbf{G}}{\left\{ E_i^{grid}(1), \cdots, E_i^{grid}(t), \cdots, E_i^{grid}(|\mathcal{T}|) \right\}} . \\ &\text{s.t.} : (5.27), (5.28), (5.29), (5.30) \end{aligned} \tag{5.36}$$

5.3.3 Optimal Offline Algorithm

In this section, the two-dimensional optimisation is studied in offline setup with non-causal information, where both amount of harvested green energy and traffic of all time slots are known in advance. The optimal offline algorithm is proposed to solve the lexicographic minimisation of on-grid energy consumption **P1** with low computational complexity when the battery capacity is infinite ($B_i^{\max} = \infty$). Since original optimisation problem **P1** is decomposed into two sub-optimisation problems, **P1** is resolved by solving two sub-optimisation problems. The solution of the user association optimisation in space dimension estimates the traffic load, thereby calculating the energy consumption of all BSs in each time slot according to equation (5.5). Based on this solution, optimal green energy allocation across different time slots is achieved by solving optimisation problem in time dimension. Hence the proposed optimal offline algorithm consists of the user association algorithm in space dimension and the green energy allocation algorithm in time dimension. User association algorithm is implemented at all BSs $\mathcal{B}$ and users at an individual time slot to determine the optimal traffic load of BSs in order to minimise total energy consumption throughout the whole network. Then based on the optimal traffic load in every time slot, green energy allocation algorithm is implemented in an individual BS to determine the green energy allocation across different time slots $\mathcal{T}$,

aiming to lexicographically minimise the on-grid energy consumption.

5.3.3.1 User Association Algorithm in Space Dimension

In this subsection, the user association algorithm is proposed to achieve the global optimum in minimising the total energy consumption of all BSs in a certain time slot. Since the user association focuses the optimisation of space dimension in a snapshot regardless of the time dimension, time slot index t is omitted in the presentation of the proposed user association algorithm for simplicity. The principle of the proposed user association in this section is similar to the algorithm in Section 5.2.3 in a sense that the user association algorithm is implemented in an iterative manner: BSs periodically measure and advertise their loads, and then users make user association decisions based on the advertised information to minimise $f(\boldsymbol{\rho}) = \sum_{i\in\mathcal{B}}[E_i + L_i(\rho_i)]$. BS and user sides update iteratively until convergence.

In order to guarantee convergence, the user arrival and departure process is assumed to be faster relative to the period in which BSs advertise their loads. Particularly, once BSs advertise their loads, users are able to make user association decisions based on advertised BS loads, prior to the next BS advertising load update.

The following elaborates the procedures of the proposed user association algorithm.

User side: At the beginning of k-th iteration, users get the traffic loads $\boldsymbol{\rho}^{(k)}$ of all BSs via broadcast. And then the user at location x chooses the optimal BS by

$$j^{(k)}(x) = \arg\max_{i\in\mathcal{B}} r_i(x)\left(\Delta_i p_i \tau + {L_i}'\left(\rho_i^{(k)}\right)\right)^{-1}. \tag{5.37}$$

where ${L_i}'\left(\rho_i^{(k)}\right) = \partial L_i\left(\rho_i^{(k)}\right)\Big/\partial\rho_i^{(k)}$. Then the user association indicator is updated by

$${y_i}^{(k)}(x) = \begin{cases} 1, & \text{if } i = j^{(k)}(x) \\ 0, & \text{otherwise,} \end{cases} \tag{5.38}$$

where $y_i^{(k)}(x)$ is broadcasted to all BSs.

BS side: The updated user association indicators from the user side will change loads of BSs, and thus during the k-th iteration, the new traffic load of BS i is given by

$$T_i\left({\rho_i}^{(k)}\right) = \min\left(\int_{\mathcal{L}} \frac{\lambda(x)\,\mu(x)\,{y_i}^{(k)}(x)}{r_i(x)} dx,\ 1-\varepsilon\right). \tag{5.39}$$

Based on the derived $T_i\left({\rho_i}^{(k)}\right)$, BS i updates the next advertising traffic load as [KdVYV12]

$${\rho_i}^{(k+1)} = \theta {\rho_i}^{(k)} + (1-\theta)\, T_i\left({\rho_i}^{(k)}\right), \quad \forall i \in \mathcal{B}, \tag{5.40}$$

where $0 \leq \theta < 1$ is an exponential averaging parameter.

The proof on optimality and convergence of the proposed user association algorithm in space dimension can be achieved with minor modifications of **Lemma 1, Lemma 2, Theorem 1,** and **Theorem 2** in Section 5.2.3.

Note that, so far, the two-dimensional optimisation considerers the condition where the **P1.1** given by (5.34) is feasible, that is the feasible set $\tilde{\mathcal{F}}$ is not empty with $\rho_i \leq 1-\varepsilon$, $\forall i \in \mathcal{B}$. However in the circumstance when the **P1.1** is not feasible due to high traffic load, the admission control is required. In this case, the admission control with the objective to minimise the sum of energy consumption and cost of blocking traffic can be formulated with the similar approach in Section 5.2.4, where a threshold is used to determine whether a particular user should be blocked or not, and the user association algorithm stays intact as the algorithm presented above in this section.

It is worthwhile mentioning that the proposed user association algorithm does not restrict the timely response for two reasons. First, the proposed user association algorithm is totally distributed. Although the interaction between users and BSs may incur overhead on control information exchange, the proposed algorithm does not require any centralised computation. As such, there is no high algorithmic complexity issue here.

Second, the duration of one time slot depends on dynamics of green energy generation and mobile traffic. As for the green energy generation, taking solar energy as an example, the granularity for solar energy generation is usually an hour [NRE]. For the mobile traffic, the hourly mobile traffic profile can well represent the traffic characteristic for guiding BS operations [PLL+11]. Thus, the time slot duration could be tens of minutes, which is long enough to execute the proposed user association algorithm. As a result, it comes to the conclusion that the proposed user association algorithm is able to be implemented whenever a new traffic request arrives and attempts the network access. Additionally, due to the constant traffic pattern during one time slot, the total energy consumption will remain the same during one time slot, regardless of the arrival and departure of the specific traffic flow.

5.3.3.2 Green Energy Allocation Algorithm in Time Dimension

Based on the non-causal traffic information, the user association algorithm in space dimension can estimate traffic load, thereby calculating the energy consumption of all BSs in each time slot according to equation (5.5). With the knowledge of energy consumption of all BSs, the green energy allocation algorithm is proposed for the green energy allocation optimisation in time dimension to lexicographically minimise the on-grid energy consumption in an offline manner, with the aid of non-causal green energy information. The green energy allocation optimisation is quite complicated, due to the fact that green energy harvested at one time slot cannot be used at its previous time slots, and the available amount of green energy during a certain time slot depends both on the green energy harvested at the current time slot and on the residual green energy harvested from previous time slots. Inspired by the energy allocation algorithm in [HA13], a green energy allocation algorithm is proposed as shown in Algorithm I.

In the proposed green energy allocation algorithm, firstly the green energy allocation

Algorithm I: **Green Energy Allocation Algorithm**

Input: $E_i(t)$, $Q_i(t), \forall t$, B_i^0;
Initialise $G_i(t), \forall t$, and calculate $E_i^{grid}(t), \forall t$;
for $m = 2; m \leq |\mathcal{T}|; m++$; **do**
  if $E_i^{grid}(m) > E_i^{grid}(m-1)$ **then**
    for $n = 1; n \leq m-1; n++$; **do**
      calculate $\bar{g} = \sum_{t=n}^{m} E_i^{grid}(t) \Big/ (m-n+1)$;
      if $E_i^{grid}(n) < \bar{g}$ **then**
        $k = n$, and break;
      end if
    end for
    for $n = k; n \leq m; n++$; **do**
      if $E_i^{grid}(n) < \bar{g}$ **then**
        Decrease $G_i(n)$ to make $E_i^{grid}(n) = \bar{g}$;
      else
        Increase $G_i(n)$ to make $E_i^{grid}(n) = \bar{g}$;
      end if
    end for
  end if
end for
Return $G_i(t), \forall t$.

is initialized as

$$G_i(t) = \begin{cases} B_i^0 + Q_i(t), & t = 1 \\ Q_i(t), & t > 1, \end{cases} \tag{5.41}$$

where BS consumes all available green energy in each time slot. Since the green energy cannot be consumed until it is harvested, the proposed green energy allocation algorithm optimises the green energy allocation of each time slot according to the time sequence. The proposed algorithm computes the on-grid energy consumption of first time slot, and then adds next time slot into green energy allocation optimisation iteratively. If the on-grid energy consumption of the newly added time slot is larger than that of the prior time slot, the green energy allocation in previous time slots will be reduced, and the saved energy will be allocated to the current time slot.

Theorem 3: The proposed green energy allocation algorithm achieves the optimal solution for **P1.2** given by (5.36), with $B_i^{\max} = \infty$.

Proof. Since the green energy harvested at one time slot cannot be used in previous time slots, the on-grid energy consumption at one time slot can be reduced only by changing

the green energy allocation in previous time slots. If $E_i^{grid}(m) > E_i^{grid}(m-1)$, the green energy allocations in time slots previous to m-th time slot are reduced, and more green energy is allocated in m-th time slot, thereby ensuring $E_i^{grid}(m) \leq E_i^{grid}(m-1)$. Specifically, we find the n-th time slot where $E_i^{grid}(n) < \sum_{t=n}^{m} E_i^{grid}(t) \Big/ (m-n+1)$, and then let $\bar{g} = \sum_{t=n}^{m} E_i^{grid}(t) \Big/ (m-n+1)$. The green energy allocation from n-th to $(m-1)$-th time slot will be reduced, and the green energy allocation in m-th slot will be increased, making sure $E_i^{grid}(t) = \bar{g}, t \in \{n, \cdots, m\}$. Assuming $E_i^{grid}(m)$ is the m-th largest on-grid energy consumption among all slots, any attempt to reduce $E_i^{grid}(m)$ will result in further increase of the largest to the $(m-1)$-th largest on-grid energy consumption. Therefore, the proposed green energy allocation achieves the min-max fair. According to [RLB07], the min-max fair vector is the unique optimal solution for the lexicographic minimisation problem, which is also Pareto optimal. □

The computational complexity of the proposed green energy allocation algorithm is $\mathcal{O}\left(|\mathcal{T}|^2\right)$ in the worst case. Such computational complexity is acceptable, since the proposed green energy allocation algorithm optimises the green energy allocation in an offline manner. If one day is taken as the whole duration of the time, the proposed green energy allocation algorithm executes only once a day.

5.3.3.3 Proof of Optimality

Theorem 4: The proposed optimal offline algorithm achieves the optimum of the original optimisation problem **P1** given by (5.26)-(5.31), with $B_i^{\max} = \infty$ and $\omega_i \to 0$.

Proof. The proposed optimal offline algorithm consists of two sequential algorithms: the user association algorithm in space dimension and the green energy allocation algorithm in time dimension. User association is first implemented to minimise the total energy consumption $\sum_{i \in \mathcal{B}} E_i(t)$ in every time slot, and let $E^*(t) = \sum_{i \in \mathcal{B}} E_i(t)$ be the resulting minimum total energy consumption in t-th time slot. Then the green energy is allocated across different time slots to lexicographically minimise the on-grid energy consumption

$\left\{E_i^{grid}(1), \cdots, E_i^{grid}(t), \cdots, E_i^{grid}(|\mathcal{T}|)\right\}$ for any individual BS, and the resulting optimal on-grid energy consumption is denoted as $\mathbf{E}^{*grid} = \left\{E_i^{*grid}(t) \,|\forall i, \forall t\right\}$. According to Algorithm I, $\mathbf{E}^{*grid}$ has the property of $E^{*grid}_i(t) \leq E^{*grid}_i(t-1), \forall t > 1$, and achieves the min-max fair in on-grid energy consumption. It is assumed that in the t-th time slot, another user association pattern with the resulting total energy consumption $\ddot{E}(t)$ is adopted, and the user association pattern in the other time slots is the same as that in the proposed optimal offline algorithm. It is obvious that $\ddot{E}(t) > E^*(t)$. If the case with $\ddot{E}(t)$ wants to achieve the same on-grid energy consumption in t-th time slot as the case in the proposed optimal offline algorithm, more green energy need to be allocated in t-th time slot. In doing so, more green energy need to be transferred from previous time slots, which may increase the on-grid energy consumption in the previous time slots. If $\ddot{\mathbf{E}}^{grid} = \left\{\ddot{E}_i^{grid}(t) \,|\forall i, \forall t\right\}$ is denoted as the resulting on-grid energy consumption in the case with $\ddot{E}(t)$, it comes to the conclusion that $\ddot{\mathbf{E}}^{grid}$ is lexicographically larger than $\mathbf{E}^{*grid}$. Therefore the proposed optimal offline algorithm, in which user association is determined to minimise the total energy consumption, and then green energy is allocated to lexicographically minimise the on-grid energy consumption, achieves the optimum of the original optimisation problem **P1** with $B_i^{\max} = \infty$ and $\omega_i \to 0$. □

Note that the proposed optimal offline algorithm provides the optimal solution based on the assumption that the traffic pattern and green energy generation are constant during one time slot. However, in real networks, the network condition may be more complicated and dynamic: the traffic pattern and green energy generation may present slight fluctuation even within a time slot. In this context, the proposed algorithm asymptotically approaches the optimal solution.

In summary, in the scenario with infinite battery capacity, the optimal offline algorithm is executed in two stages. First, user association is determined to minimise the total energy consumption in every time slot. Then green energy is allocated across different time slots to lexicographically minimise the on-grid energy consumption. However, in the general case, if the battery capacity is finite, the green energy cannot be arbitrarily

saved for future use due to the limited battery capacity. The structure of optimal green energy allocation cannot be presented in a simple and clear way as Algorithm I in Section 5.3.3.2. Intuitively, in contrast with the case of infinite battery capacity where green energy is conservatively used to achieve optimality, in the case of finite battery capacity the green energy should be allocated first to minimise the energy overflow. Some heuristic online algorithms will be proposed based on this intuition, and the proposed optimal offline algorithm in the infinite battery capacity can be treated as the performance upper bound of online algorithms with finite battery capacity.

5.3.4 Heuristic Online Algorithms

In this section, the formulated two-dimensional optimisation problem **P1** given by (5.26)-(5.31) is investigated in online setup with causal information, where only the amount of harvested green energy and traffic in the current and previous time slots are known. Furthermore, the historical green energy and mobile traffic statistics are available to estimate the harvested green energy and traffic across different time slots [NRE, San].

The user association optimisation in the proposed optimal offline algorithm in Section 5.3.3.1 decides the user association in a certain time slot to minimise the total energy consumption. As mentioned in Section 5.3.3.1, the proposed user association algorithm is capable of timely response and only requires the traffic information of the current time slot. In this sense, it can already be implemented in the online manner. The green energy allocation in the proposed optimal offline algorithm in Section 5.3.3.2 requires the non-causal green energy and traffic knowledge, which is not practical. But its structure sheds light on the design of online algorithms. As such, in this section, heuristic online green energy allocation algorithms are proposed with online setup. More specifically, the online green energy allocation will update in each time slot according to the amount of energy consumption and available green energy, as well as the battery capacity.

5.3.4.1 Constant On-Grid Energy Consumption Level Algorithm

Motivated by the green energy allocation in the proposed optimal offline algorithm presented in Section 5.3.3.2 which tries to achieve uniform on-grid energy consumption as much as possible, a constant on-grid energy consumption level algorithm is proposed. The constant on-grid energy consumption level is defined as

$$\tilde{g}_i = \left(\sum_{t \in \mathcal{T}} \tilde{E}_i(t) - \sum_{t \in \mathcal{T}} \tilde{Q}_i(t) - B_i^0 \right) \Big/ |\mathcal{T}|, \tag{5.42}$$

where $\tilde{E}_i(t)$ is statistics on total energy consumption of BS i at t-th time slot, which can be calculated based on the historical traffic statistics and the proposed user association algorithm presented in Section 5.3.3.1. $\tilde{Q}_i(t)$ is the historical statistics on harvested green energy of BS i at t-th time slot. Then the intended green energy allocation of BS i at t-th time slot is

$$\tilde{G}_i(t) = \max\left\{ E_i(t) - \tilde{g}_i, 0 \right\}. \tag{5.43}$$

However, the allocated green energy cannot be larger than the maximum available green energy of BS i at t-th time slot, which is given by

$$G_i^{\max}(t) = R_i(t) + Q_i(t). \tag{5.44}$$

Thus the green energy allocation of BS i at t-th time slot is calculated as

$$G_i(t) = \min\left\{ \tilde{G}_i(t), G_i^{\max}(t) \right\}. \tag{5.45}$$

As such, the on-grid energy consumption of BS i at t-th time slot is

$$E_i^{grid}(t) = \max\left\{ E_i(t) - G_i(t), 0 \right\}. \tag{5.46}$$

Due to the limit on battery capacity and the randomness of green energy arrival, green energy overflow in the online setup may exist. Hence in addition, the *green energy overflow prevention* is considered. If the expected stored energy exceeds the battery capacity $B_i^{\max}$, the minimum green energy allocation is

$$G_i^{\min}(t) = R_i(t) + Q_i(t) - B_i^{\max}. \tag{5.47}$$

As such, the green energy allocation of BS i at time slot t is determined as

$$G_i(t) = \min\left\{\max\left\{\tilde{G}_i(t), G_i^{\min}(t)\right\}, G_i^{\max}(t)\right\}. \tag{5.48}$$

The green energy overflow prevention well meets the battery capacity constraint from the average point of view, and is expected to improve the performance, although the green energy overflow can not be completely avoided.

5.3.4.2 Adaptive On-Grid Energy Consumption Level Algorithm

For finite time transmissions, the constant on-grid energy consumption level algorithm is apparently not optimal due to the variation of traffic load and available green energy across different time slots. Thus the adaptive on-grid energy consumption level algorithm is proposed to improve the performance. The on-grid energy consumption level is updated for each time slot, which is given by

$$\tilde{g}_i(t) = \frac{E_i(t) + \sum_{m=t+1}^{|\mathcal{T}|}\tilde{E}_i(m) - \left(R_i(t) + Q_i(t) + \sum_{m=t+1}^{|\mathcal{T}|}\tilde{Q}_i(m)\right)}{|\mathcal{T}| - t + 1}. \tag{5.49}$$

In this case, the intended green energy allocation of BS i at t-th time slot is

$$\tilde{G}_i(t) = \max\{E_i(t) - \tilde{g}_i(t), 0\}. \tag{5.50}$$

Then the green energy allocation in the adaptive on-grid energy consumption level algorithm without and with *green energy overflow prevention* are presented as same as (5.45) and (5.48), respectively.

5.3.5 Simulation Results and Conclusions

5.3.5.1 Simulation Results

As mentioned in Section 5.3.3.1, the proposed user association algorithm is able to be executed within one time slot. Note that the theoretical analysis throughout paper is independent with the exact duration of one time slot, and also not restricted to a particular type of renewable energy source. To evaluate the effectiveness of proposed algorithms in accommodating the temporal dynamic of mobile traffic and green energy generation, the statistic in [NRE, San] is adopted in the simulation, where the statistics in [NRE] provides hourly solar energy generation in June around London gatwick airport, and [San] estimates the typical daily traffic variation in European areas as shown in Fig. 5.9. The daily traffic variation is plotted relative to the peak traffic demand. In the simulation, all BSs are assumed to be in the same weather environment and bear the same traffic variation trend. The performance of proposed algorithms is simulated during the time scale of 24 hours.

Figs. 5.10-5.14 evaluate the performance of the proposed optimal offline algorithm in the scenario with infinite battery capacity. Figs. 5.15-5.16 demonstrate the performance of proposed online algorithms with finite battery capacity. Here the battery capacity of macrocell [1] is set as S times of the average harvested green energy in one time slot,

[1] In the simulation, both energy harvesting rate and battery capacity of picocell are assumed as 2.5% of those of marcocell. This assumed ratio is reasonable, since the transmit power ratio of picocell and

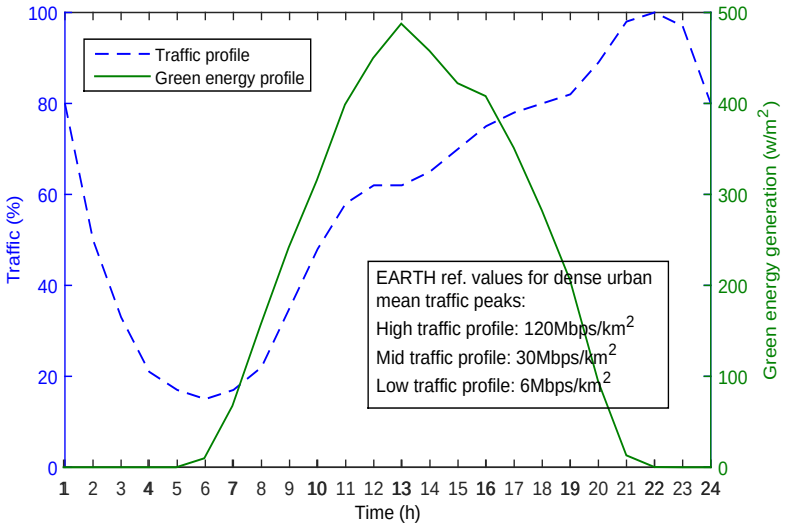
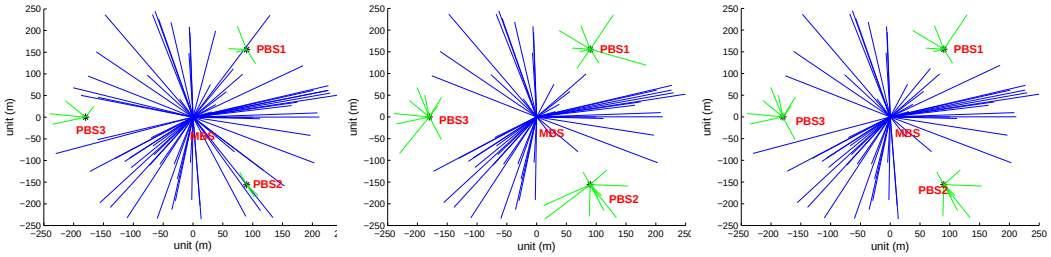


Figure 5.9: Traffic and green energy profiles versus different time slots.



(a) Without optimisation in space dimension.

(b) With optimisation but no penalty function in space dimension.

(c) With optimisation and penalty function in space dimension.

Figure 5.10: Snapshot of user association

where S is called relative battery capacity. The proposed optimal offline algorithm in the infinite battery capacity can be treated as the performance upper bound for online algorithms in finite battery capacity.

5.3.5.2 Behaviour of the Proposed Optimal Offline Algorithm

Figs. 5.10-5.11 focus on the snapshot of the time slot with peak traffic demand in the mid traffic profile condition, and verify the effectiveness of the user association algorithm in space dimension in the proposed optimal offline algorithm.

---

marcocell is 2.5% defined in the 3GPP standards.

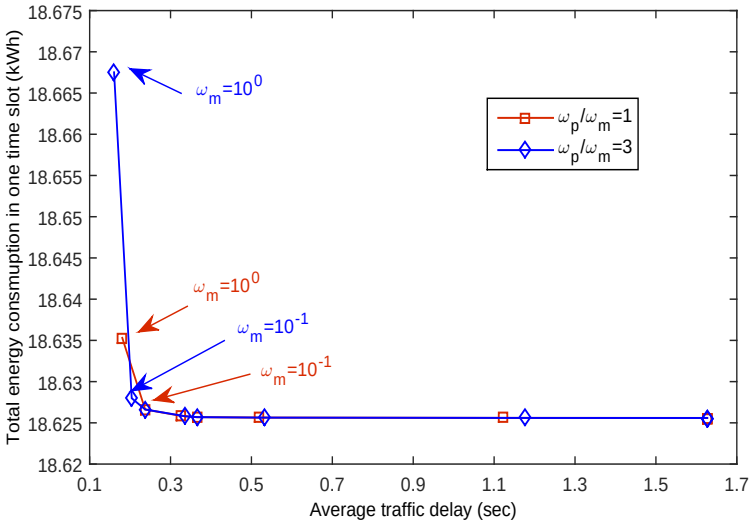


Figure 5.11: Tradeoff between average traffic delay and total energy consumption by varying the weight of macrocell from 10^0 to 10^{-20}.

Fig. 5.10 compares the snapshot of the resulting user association pattern in one macrocell area with and without the optimization/penalty function in space dimension. In the user association without optimisation, users will associate with the BS which provides the strongest DL RSS. Due to the power disparity of the high power MBS and low power PBS, Fig. 5.10 demonstrates that in the user association without optimisation most users associate with MBS, which will overload the marcocell and make the PBS deployment ineffective. On the contrary, in the user association with optimisation in space dimension without penalty function, where user association is determined by minimizing $\sum_{i\in\mathcal{B}}[E_i(t)]$, it is obvious that many users associate with PBSs. Although it effectively offloads the traffic from the congested macrocell to picocells, due to the limited capacity of picocells, excessive offloaded users may congest picocells. The proposed user association with optimisation and penalty function in space dimension compromises the above two schemes, where users are well-balanced and the probability of congestion is further reduced. Such traffic balancing achieves a good tradeoff between energy saving and average traffic delay reduction which is demonstrated in Fig. 5.11.

The overall load balancing achieved by adding the penalty function comes at the cost of slight increase in the total energy consumption. Here, the load balancing benefit is

quantised by the average traffic delay. Users associated with the same BS are assumed to be served on the round robin fashion. Considering the system as M/GI/1 multi-class processor sharing system as in [J.W98], $\rho_i/(1-\rho_i)$ is equal to the average number of flows at BS i, and $\sum_{i\in\mathcal{B}} \frac{\rho_i}{1-\rho_i}$ is the total number of flows in the system [KdVYV12]. According to the Little's law, the average number of flows is mathematically related to the average delay experienced by a typical traffic flow. Fig. 5.11 calculates average traffic delay and total energy consumption with different weights of macrocell ω_m. The less average traffic delay means the less congestion and more effective load balancing. As shown in Fig. 5.11, the increment of total energy consumption is marginal compared with the average traffic delay reduction. For example, in the case of $\omega_m = 10^0$ and $\omega_p/\omega_m = 3$, where ω_p denotes the weights of picocell, there is 90% reduction in average traffic delay with 0.22% increase in total energy consumption compared with the case without penalty function ($\omega_m \to 0$). In addition, it is shown that the performance of $\omega_p/\omega_m = 3$ is better than that of $\omega_p/\omega_m = 1$. Taking $\omega_m = 10^0$ as an example, compared with the case of $\omega_p/\omega_m = 1$, the case of $\omega_p/\omega_m = 3$ can get 11.2% more reduction in the average traffic delay, but only increases the total energy consumption by 0.17%. This verifies the effectiveness of making $\omega_p/\omega_m > 1$ for HetNets to avoid the early capacity bottleneck of picocells. Note that such tradeoff graph may also be used to choose the value of ω_m and ω_p/ω_m in practice based on the maximum tolerable traffic delay. In the following simulation of this section, ω_m and ω_p/ω_m are set as 10^0 and 3, respectively.

Fig. 5.12 testifies the effectiveness of the green energy allocation algorithm in time dimension in the proposed optimal offline algorithm. Fig. 5.12 shows the green energy allocation and on-grid energy consumption versus different time slots with and without optimisation in time dimension. In the case without optimisation, BSs consume the green energy as long as they harvest it from renewable energy sources. Fig. 5.12 illustrates that the proposed green energy allocation algorithm in time dimension optimises the green energy allocation over time slots. As a result, on-grid energy is consumed in a more uniform way with respect to time, which mitigates the high peak on-grid energy

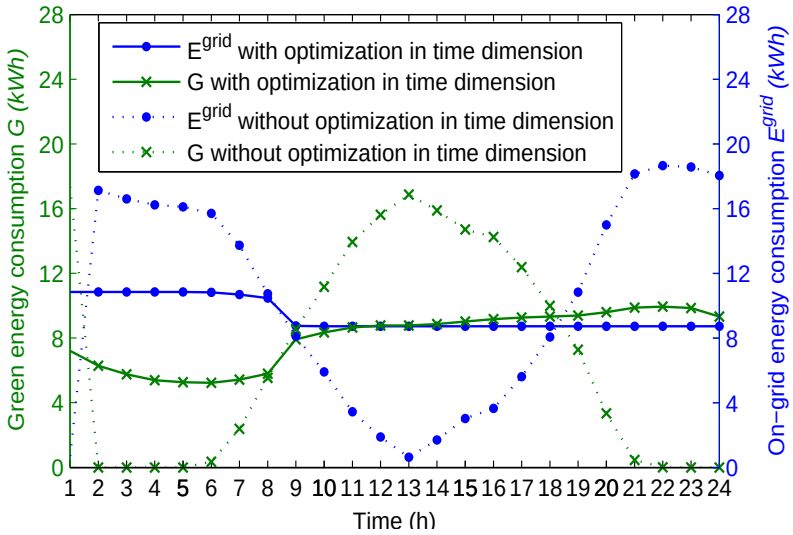


Figure 5.12: Green energy allocation and on-grid energy consumption versus different time slots o/w optimisation in time dimension.

consumption problem. However, in the case without optimisation, due to the high green energy generation rate during 10:00 and 16:00 as shown in Fig. 5.9, the on-grid energy consumption goes down dramatically, but rapidly escalates after 19:00, since the green energy generation rate experiences rapid decline after then.

5.3.5.3 Comparison of the Proposed Optimal Offline Algorithm with Baseline Algorithms

For comparison, three baseline algorithms are considered for the proposed optimal offline algorithm. In baseline algorithm 1, there is no optimisation in neither space nor time dimension. Baseline algorithm 2 and 3 only have optimisation in space and time dimension, respectively. Fig. 5.13 presents the total on-grid energy consumption within the whole HetNets area and across all time slots in different traffic profiles. Compared with the baseline algorithm 1, the baseline algorithm 2 and 3 can reduce the total on-grid energy consumption. The superiority of baseline algorithm 2 over baseline algorithm 1 demonstrates that the load balancing achieved in the proposed user association algorithm effectively benefits the total energy consumption minimisation. The proposed optimal

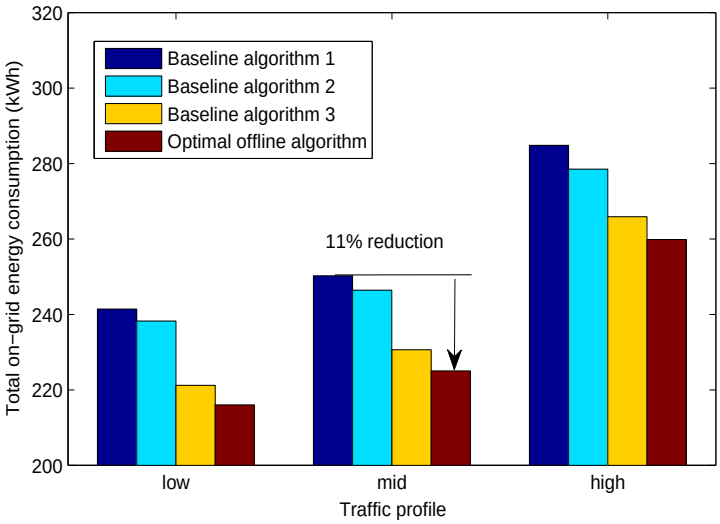


Figure 5.13: Total on-grid energy consumption in different traffic profiles. Only offline algorithms are considered.

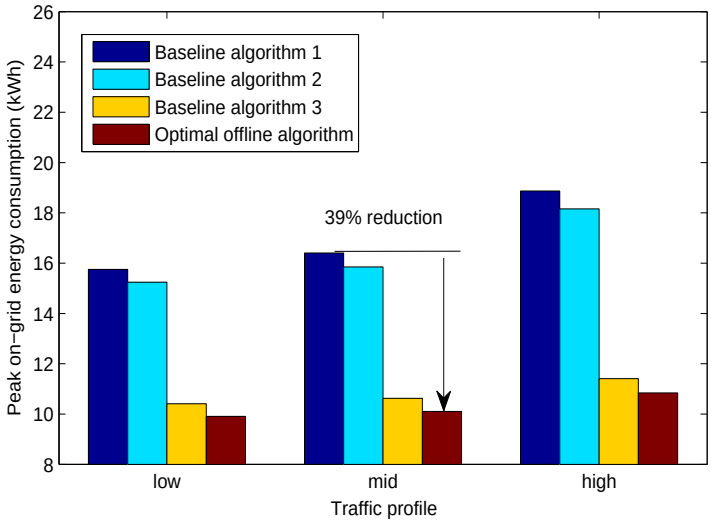


Figure 5.14: Peak on-grid energy consumption in different traffic profiles. Only offline algorithms are considered.

offline algorithm with optimisation in both time and space dimensions achieves the most on-grid energy saving among these four algorithms, with about 10% total on-grid energy saving, benchmarked by the baseline algorithm 1.

Fig. 5.14 evaluates the peak on-grid energy consumption in different traffic profiles. The peak on-grid energy consumption plotted here is the sum peak on-grid energy consumption of all BSs in one macrocell area. Fig. 5.14 reveals that compared with base-

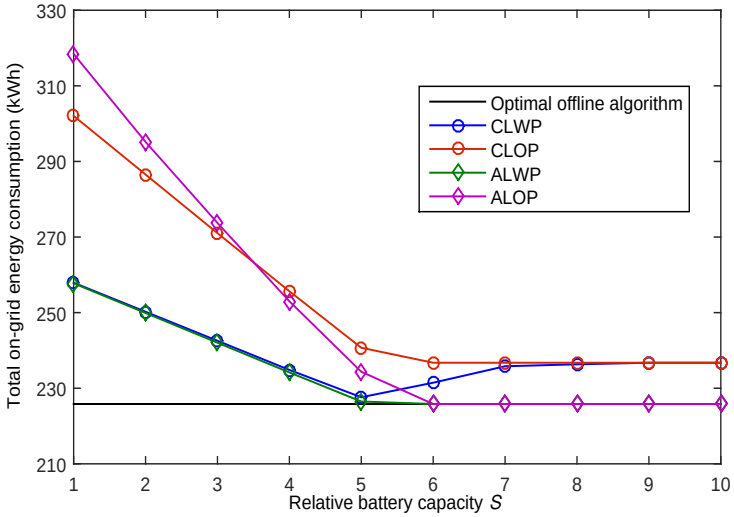


Figure 5.15: Total on-grid energy consumption versus different battery capacities.

line algorithms, the proposed optimal offline algorithm with optimisation in both time and space dimensions is able to substantially reduce the peak on-grid energy consumption, with about 40% reduction compared to the baseline algorithm 1. It indisputably addresses the challenge of rising peak on-grid energy consumption, thereby reducing OPEX and maintaining profitability for mobile network operators.

Figs. 5.13-5.14 demonstrate the design rationale of the proposed optimal offline algorithm which involves the optimisation in both time and space dimensions, since it is more effective than the algorithm with optimisation in time or space dimension only.

5.3.5.4 Comparison of Proposed online and Optimal Offline Algorithms

Fig. 5.15 shows the total on-grid energy consumption of different online algorithms versus different battery capacities. They are benchmarked by the proposed optimal offline algorithm with infinite battery capacity, which is the performance upper bound for all online algorithms. To facilitate the subsequent explanation, the proposed constant on-grid energy consumption level algorithm without and with green energy overflow prevention are denoted as CLOP and CLWP, respectively. Also, the proposed adaptive on-grid

energy consumption level algorithm without and with green energy overflow prevention are denoted as ALOP and ALWP, respectively. As shown in Fig. 5.15, in the low battery capacity regime, say $S < 3$ in this simulation, the CLOP outperforms the ALOP. This can be explained by the fact that ALOP updates the on-grid energy consumption level periodically, and foresees the low green energy generation rate and high traffic load during late night, so more green energy is intended to be saved for the late night. However since there is no green energy overflow prevention, there is a higher chance that the conserved energy will exceed the low battery capacity, resulting in energy overflow. Fig. 5.15 also reveals that in CLWP, with the growing battery capacity, the total on-grid energy consumption first decreases and then increases. This is because with the increase of battery capacity, more green energy may be saved for the subsequent time slots, which benefits the on-grid energy saving. However when the relative battery capacity exceeds a certain value, say $S > 5$ in this simulation, consuming all the conserved green energy may result in the on-grid energy consumption in the subsequent time slots smaller than the predefined constant on-grid energy consumption level. Due to the fact that the on-grid energy consumption in CLWP cannot be smaller than the predefined constant level, the excessive conserved green energy may be wasted, giving rise to the total on-grid energy consumption increase. In addition, it is concluded from Fig. 5.15 that the green energy overflow prevention is efficient in total on-grid energy saving. It substantially reduces the green energy overflow at low battery capacity regime, and the gain shrinks as battery capacity increases. It is also shown that the performance of ALOP and ALWP approaches that of the optimal offline algorithm with increasing battery capacity.

Fig. 5.16 demonstrates the effect of battery capacity on the peak on-grid energy consumption in different online algorithms. It shows that as battery capacity increases, the peak on-grid energy consumption of online algorithms reduces, and the gap between online algorithms and optimal offline algorithm decreases. It is also observed that ALOP and ALWP outperform CLOP and CLWP, since ALOP and ALWP are able to update the on-grid energy consumption level adaptively, thereby better utilising the harvested green

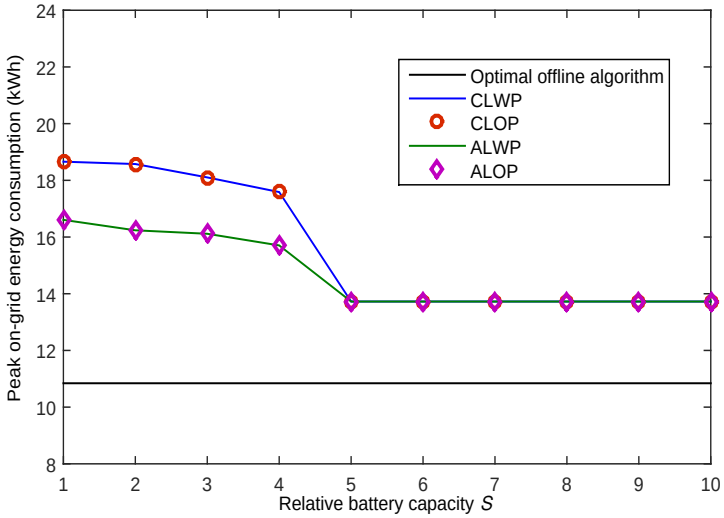


Figure 5.16: Peak on-grid energy consumption versus different battery capacities.

energy. Based on Figs. 5.15-5.16, it comes to the conclusion that ALWP is superior to the other proposed online algorithms no matter what the size of battery capacity, considering both total and peak on-grid energy consumption reductions.

5.3.5.5 Conclusions

In section 5.3, the two-dimensional optimisation on user association and green energy allocation was studied to lexicographically minimise the on-grid energy consumption in HetNets with hybrid energy sources, where all BSs are assumed to be powered by both power grid and renewable energy sources. The optimisation problem was decomposed into two sub-optimisation problems: the user association optimisation in space dimension and the green energy allocation optimisation in time dimension. The low complexity optimal offline algorithm with infinite battery capacity was first developed by assuming non-causal green energy and traffic information, which can serve as performance upper bound for evaluating practical online algorithms. Simulation results indicate the proposed optimal offline algorithm substantially saves on-grid energy as well as reduces peak on-grid energy consumption. In addition, motivated by the optimal offline algorithm,

some heuristic online algorithms with finite battery capacity, utilising causal green energy and traffic information only, were also proposed and evaluated by simulations. Simulation results demonstrate that the adaptive on-grid energy consumption level algorithm with green energy overflow prevention outperforms the other proposed heuristic online algorithms in terms of both total and peak on-grid energy consumption reductions.

5.4 Summary

This chapter focused on the user association optimisation in HetNets with hybrid energy sources, where BSs are envisioned to be powered by both power grid and renewable energy sources.

Section 5.1 specified the general system model and simulation platform employed in this chapter.

An optimal user association algorithm was developed to achieve the tradeoffs between average traffic delay and on-grid energy consumption in Section 5.2. The proposed user association algorithm allows for a flexible tradeoff between average traffic delay and on-grid energy consumption by adjusting the value of weight ω. Simulation results demonstrate that the proposed user association algorithm is able to adapt loads of BSs along with distributions of green energy, thereby effectively reducing on-grid energy consumption, as well as achieving comparable average traffic delay compared to the existing algorithm which aims to minimise the average traffic delay.

In addition, the two-dimensional optimisation on user association and green energy allocation was investigated to minimise both total and peak on-grid energy consumptions in Section 5.3. Both optimal offline and heuristic online algorithms were designed for such a scenario. Simulation results indicate that the proposed optimal offline algorithm achieves more than 10% total on-grid energy consumption reduction and about 40% peak on-grid energy consumption reduction. Furthermore, the proposed low complex-

ity optimal offline algorithm with infinite battery capacity is capable of serving as the performance upper bound for evaluating practical online algorithms. Among the four proposed heuristic online algorithms, the proposed adaptive on-grid energy consumption level algorithm with green energy overflow prevention outperforms the other proposed heuristic online algorithms in terms of both total and peak on-grid energy consumption reductions.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis was dedicated to the user association optimisation in different HetNets scenarios.

In conventional grid-powered HetNets, the idea of NBS from cooperative game theory can be efficiently applied to the user association optimisation. Compared with the state-of-the-art user association algorithms, the NBS based user association algorithms proposed in Chapter 3 work better in addressing load balancing, user fairness, spectrum efficiency, energy efficiency, UL-DL asymmetry and diverse QoS provision issues in HetNets.

In HetNets with renewable energy powered BSs, the optimal offline and heuristic online algorithms were proposed for adaptive user association in Chapter 4. As one of the pioneer works on user association optimisation in HetNets with renewable energy powered BSs, the proposed algorithms are able to adjust the user association decision according to renewable energy and load variations, as well as guarantee the QoS provision for users.

The HetNets with hybrid energy sources, where BSs powered by both power grid and renewable energy sources, have the superiority in supporting uninterrupted service as

well as achieving green communications. In this context, the optimal user association algorithm proposed in Section 5.2 is capable of achieving the tradeoffs between average traffic delay and on-grid energy consumption. The two-dimensional optimisation on user association and green energy allocation proposed in Section 5.3 is able to enhance the QoS provision, as well as substantially reduce both total and peak on-grid energy consumptions, thereby indisputably reducing OPEX and maintaining profitability for mobile network operators.

For all algorithms proposed in this thesis, great attention is given to accommodate the inherent nature of HetNets and renewable energy sources in the user association algorithm design. The proposed algorithms provide useful guidelines and potential solutions for the user association mechanisms in future HetNets.

6.2 Future Work

6.2.1 User Association in Energy Cooperation Enabled Networks

The deployment of renewable energy sources to supplement conventional power grid for powering BSs indisputably underpins the trend of green communication. In addition, the smart grid, as one of user cases envisioned for 5G networks [Nok14], has paved the way for energy cooperation in networks. Energy cooperation between BSs allows the BSs that have excessive harvested renewable energy to compensate for others that have a deficit with the aid of renewable energy transfer. User association in energy cooperation enabled networks introduces the tradeoff between signal strength degradation caused by traffic offloading and renewable energy loss caused by energy transfer, which is a very challenging research topic. To the best of our knowledge, user association in energy cooperation enabled networks is still a fairly open field, and is expected to become a rewarding research area.

6.2.2 User Association in Massive MIMO Enabled HetNets

HetNets escalates the spectrum efficiency by using a mix of macrocells and small cells. On the other hand, massive MIMO technology promises enormous enhancement in spectrum efficiency by transmitting independent data streams via a tremendous number of antennas simultaneously to multiple users sharing the same transmission resource. These benefits have put massive MIMO and HetNets in the spotlight of preliminary 5G discussions [ABC+14].

The implementation of massive MIMO has a big effect on user association. The research on user association in massive MIMO enabled HetNets is limited and still in its infancy. User association schemes designed for the existing cellular system may not be capable of addressing the effect of large multiplexing and array gains provided by massive MIMO BS. For the upcoming HetNets employing massive MIMO for 5G, the design of new user association schemes is indispensable, and there are at least two aspects that should be taken into account: 1) max RSS based user association may force the massive MIMO BS to carry the most of data traffic in HetNets, due to the large array gain achieved by massive MIMO. Therefore, throughput load balancing is important in massive MIMO enabled HetNets; 2) Although massive MIMO uses large numbers of antennas and requires more power for complex signal processing, it still can be energy efficient by serving more users, since the power consumption per user is reduced and more spectrum efficiency is achieved. As such, energy efficient user association in massive MIMO enabled HetNets should address the interplay between the number of antennas at the BS and number of users served by the massive MIMO BS.

Therefore, it comes to the conclusion that user association in massive MIMO enabled HetNets is a promising research avenue, and more research efforts are needed for the final practical deployment of 5G networks.

Appendix A

Verification and Validation

The work in this thesis is a theoretical analysis and optimisation. All derivations are based on the assumptions from related 3GPP standards and fundamental equations in communications theory, which are clearly stated accordingly in the thesis. The correctness of derivations can be checked by tracing back all substitutions and manipulations through mathematical calculations.

In order to testify the performance of the proposed algorithms, the theoretical framework is translated to the Matlab code. In the simulation, the typical parameter values specified in related 3GPP standards and previous work are adopted, as shown in the corresponding references in Section 3.1.2, Section 4.6.1 and Section 5.1.2. All codes are checked by the debugger, and the mathematical correctness of the implemented algorithm is checked line by line by hand.

To verify the correctness of my coding, reproduction of the related previous work is carried out. For instance, in Chapter 3, Fig. 3.4 shows that the max RSS algorithm will result in about 10% users being associated with PBSs if there are 3 PBSs in HetNets area, which agrees with the result in [Guv11, Fig. 3]. Fig. 3.4 shows that more than 80% users are associated with PBSs in max sum rate algorithm proposed in [CFM12], which matches the original result in [CFM12]. In chapter 5, Fig. 5.10 demonstrates the effect of penalty function on offloading traffic from congested cells, which is in line with the effect of similar penalty function in [SNSD13].

To evaluate the new algorithms proposed in this thesis, the performance of proposed algorithms is compared with existing algorithms, thereby ensuring the trend of the simulation results is reasonable. For example, in Chapter 3, Fig. 3.7 indicates that the proposed scheme with Hungarian algorithm converges in about one to six rounds, and this is in-line with the results in the previous work [HJL05], which also adopts the similar Hungarian algorithm. In Chapter 5, Fig. 5.6 indicates that both in the proposed and the existing algorithm in [KdVYV12], the average traffic delay increases with the increment of traffic arrival rate, and this trend agrees with the result in [KdVYV12].

Validation of work is carried out step by step in each stage. This reduces the chance of error at a later stage of development and provides a return path if there is any mistake.

Appendix B

Simulation Drops Justification

The location of BSs is assumed fixed in all simulations in this thesis. Nevertheless, users are randomly distributed in the simulation in Chapter 3 and Chapter 4, and file transfer requests are simulated to follow a homogenous Poisson point process where $\lambda(x) = \lambda$ in Chapter 5. As such, the location of users changes in each simulation drop in Chapter 3 and Chapter 4, and both the number of traffic requests and their location vary in each simulation drop in Chapter 5. In order to average out the effect of such random factors and obtain reliable index of performance indicators, Monte Carlo simulations are adopted in the performance evaluation. However, the number of drop used in Monte Carlo simulations makes a lot of difference to the reliability and the feasibility of simulation results. Excessive drops will consume a plethora of computer resources and take long time to obtain simulation results, whereas insufficient drops will lead to unreliable results. In order to choose a reasonable number of simulation drops, the values of blocking probability in Fig. 5.8 with $\alpha^{-1} = 5000$ are compared under different simulation drops as shown in Fig. A.1 and Fig. A.2.

Fig. A.1 shows the PDFs of blocking probability under different simulation drops. It is obvious that the standard deviation (SD) under 500 simulation drops is much smaller than that under 50, 100, 200 drops, which indicates the result under 500 simulation

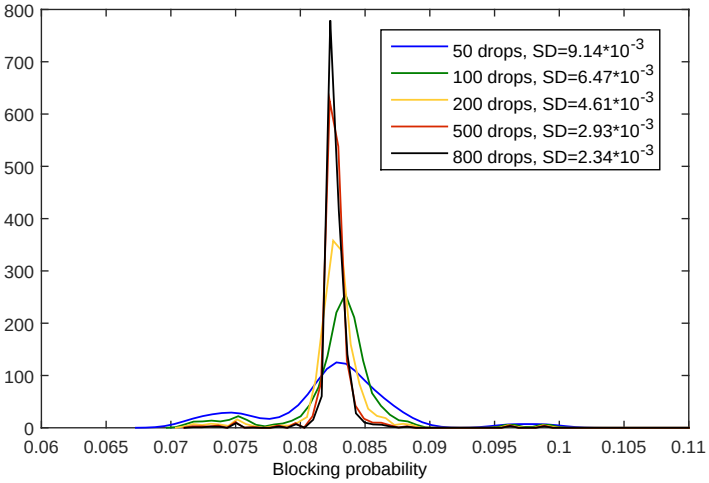


Figure A.1: PDFs of blocking probability under different simulation drops.

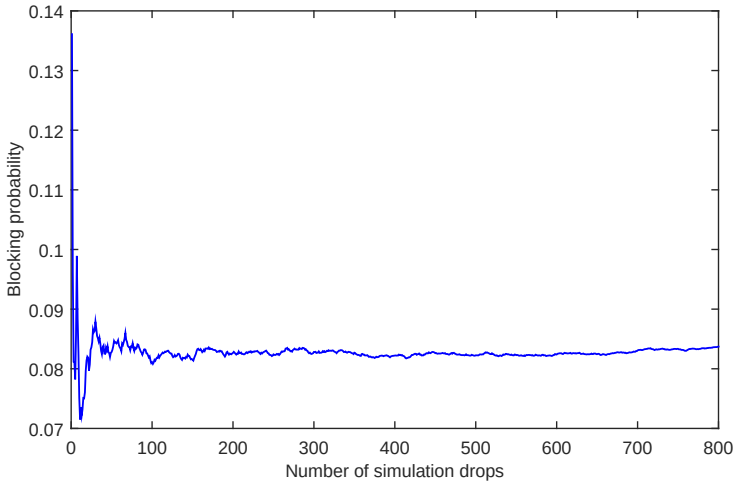


Figure A.2: Blocking probability versus different numbers of simulation drops.

drops is comparatively more accurate. In addition, the SD under 500 simulation drops is comparable to the SD under 800 drops.

Fig. A.2 demonstrates the blocking probability versus different numbers of simulation drops. The fluctuation of blocking probability is negligible after 500 simulation drops. As such, the number of simulation drops is set as 500 in the simulation in Chapter 5.

The number of simulation drops in Chapter 3 and Chapter 4 can be justified in a similar method.

References

[3GP02] 3GPP. TS 23.107 quality of service (QoS) concept and architecture. *3GPP TSG RAN WG1 Meeting-62*, Mar. 2002.

[3GP10] 3GPP. Further advancements for E-UTRA physical layer aspects (TR 36.814). Mar. 2010.

[3gp15] Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); overall description: Stage 2, 3GPP TS 36.300 v12.5.0, release 12. Apr. 2015.

[5gp] 5G Infrasrucute Public Private Partnership. `Available:http://5g-ppp.eu/`.

[ABC+14] J.G. Andrews, S. Buzzi, Wan Choi, S.V. Hanly, A. Lozano, A.C.K. Soong, and J.C. Zhang. What will 5G be? *IEEE J. Sel. Areas Commun.*, 32(6):1065–1082, Jun. 2014.

[AGD+11] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M.A. Imran, D. Sabella, M.J. Gonzalez, O. Blume, and A. Fehske. How much energy is needed to run a wireless network? *IEEE Wireless Commun.*, 18(5):40–49, Oct. 2011.

[BJ07] A.V. Babu and L. Jacob. Fairness analysis of IEEE 802.11 multirate wireless LANs. *Vehicular Technology, IEEE Transactions on*, 56(5):3073–3088, Sep. 2007.

[BLM+14] N. Bhushan, Junyi Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer. Network densification: the dominant theme for wireless evolution into 5G. *IEEE Commun. Mag.*, 52(2):82–89, Feb. 2014.

[CAG08] V. Chandrasekhar, J.G. Andrews, and Alan Gatherer. Femtocell networks: a survey. *IEEE Commun. Mag.*, 46(9):59–67, Sep. 2008.

[CFM12] S. Corroy, L. Falconetti, and R. Mathar. Dynamic cell association for downlink sum rate maximization in multi-cell heterogeneous networks. In *Proc. IEEE Int. Conf. on Commun. (ICC)*, pages 2457–2461, Jun.

2012.

[CH12] Xue Chen and R.Q. Hu. Joint uplink and downlink optimal mobile association in a wireless heterogeneous network. In *Proc. 2012 IEEE Global Commun. Conf. (GLOBECOM)*, pages 4131–4137, Dec. 2012.

[Cis15] Cisco. Cisco visual networking index: Global mobile data traffic forecast update 2014-2019 white paper. Feb. 2015.

[CLW12] Ying Cui, V.K.N. Lau, and Yueping Wu. Delay-aware BS discontinuous transmission control and user scheduling for energy harvesting downlink coordinated MIMO systems. *IEEE Trans. Signal Process.*, 60(7):3786–3795, Jul. 2012.

[CRAF15] E. Chavarria Reyes, I. Akyildiz, and E. Fadel. Energy consumption analysis and minimization in multi-layer heterogeneous wireless systems. *IEEE Trans. Mobile Computing*, PP(99):1–1, 2015.

[CZXL11] Yan Chen, Shunqing Zhang, Shugong Xu, and G.Y. Li. Fundamental trade-offs on green wireless networks. *IEEE Commun. Mag.*, 49(6):30–37, Jun. 2011.

[Dep14] United States Department of Energy. Peak-to-average electricity demand ratio rising in new england and many other U.S. regions. *U.S. Energy Information Administration Independent Statistic and Analysis*, Feb. 2014.

[DLN+14] H.S. Dhillon, Ying Li, P. Nuggehalli, Zhouyue Pi, and J.G. Andrews. Fundamentals of heterogeneous cellular networks with energy harvesting. *IEEE Trans. on Wireless Commun.*, 13(5):2782–2797, May 2014.

[DMMS14] S. Deb, P. Monogioudis, J. Miernik, and J.P. Seymour. Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets. *IEEE/ACM Trans. Networking*, 22(1):137–150, Feb. 2014.

[DMW+11] A. Damnjanovic, J. Montojo, Yongbin Wei, Tingfang Ji, Tao Luo, M. Vajapeyam, Taesang Yoo, Osok Song, and D. Malladi. A survey on 3GPP heterogeneous networks. *IEEE Wireless Commun.*, 18(3):10–21, Jun. 2011.

[DOC10] NTT DOCOMO. Performance of eICIC with control channel coverage limitation. *3GPP TSG RAN WG1 Meeting 61,R1-103264*, May 2010.

[Eri12] Ericsson. Towards heterogeneous networks. `Available:http://www.cwc.oulu.fi/researchseminar2012/Handouts/Ericsson.pdf`, FEB. 2012.

[FADR11] D. Fooladivanda, A. Al Daoud, and C. Rosenberg. Joint channel allocation and user association for heterogeneous wireless cellular networks. In *Proc. 2011 IEEE 22nd Int. Sym. on Personal Indoor and Mobile Radio Commun. (PIMRC)*, pages 384–390, Sep. 2011.

[FJL+13] Daquan Feng, Chenzi Jiang, Gubong Lim, Jr. Cimini, L.J., Gang Feng, and G.Y. Li. A survey of energy-efficient wireless communications. *IEEE Commun. Surveys Tutorials*, 15(1):167–178, First 2013.

[GR13] Jagadish Ghimire and Catherine Rosenberg. Resource allocation, transmission coordination and user association in heterogeneous networks: A flow-based unified approach. *IEEE Trans. Wireless Commun.*, 12(3):1340–1351, Mar. 2013.

[GSSBH11] A. Garcia-Saavedra, P. Serrano, A. Banchs, and M. Hollick. Energy-efficient fair channel access for IEEE 802.11 WLANs. In *Proc. 2011 IEEE Int. Sym. on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–9, Jun. 2011.

[Guv11] I. Guvenc. Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination. *IEEE Commun. Lett.*, 15(10):1084–1087, Oct. 2011.

[GWZ13] M. Gorlatova, A Wallwater, and G. Zussman. Networking low-power energy harvesting devices: Measurements and algorithms. *IEEE Trans. Mobile Comput.*, 12(9):1853–1865, Sept. 2013.

[GZN13] Jie Gong, Sheng Zhou, and Zhisheng Niu. Optimal power allocation for energy harvesting and power grid coexisting wireless communication systems. *IEEE Trans. Commun.*, 61(7):3040–3049, Jul. 2013.

[HA12] Tao Han and N. Ansari. Ice: Intelligent cell breathing to optimize the utilization of green energy. *IEEE Commun. Lett.*, 16(6):866–869, Jun. 2012.

[HA13] Tao Han and N. Ansari. On optimizing green energy utilization for

cellular networks with hybrid energy supplies. *IEEE Trans. Wireless Commun.*, 12(8):3872–3882, Aug. 2013.

[HH15] E. Hossain and M. Hasan. 5G cellular: key enabling technologies and research challenges. *IEEE Instrumentation Measurement Mag.*, 18(3):11–21, Jun. 2015.

[HJL05] Zhu Han, Z. Ji, and K.J.R. Liu. Fair multiuser channel allocation for OFDMA networks using nash bargaining solutions and coalitions. *IEEE Trans. Commun.*, 53(8):1366–1376, Aug. 2005.

[HL13] Mingyi Hong and Zhi-Quan Luo. Distributed linear precoder optimization and base station selection for an uplink heterogeneous network. *IEEE Trans. Signal Process.*, 61(12):3214–3228, Jun. 2013.

[HL14] Vu Nguyen Ha and Long Bao Le. Distributed base station association and power control for heterogeneous cellular networks. *IEEE Trans. Veh. Tech.*, 63(1):282–296, Jan. 2014.

[HM12] J. Hoadley and P. Maveddat. Enabling small cell deployment with HetNet. *IEEE Wireless Commun.*, 19(2):4–5, Apr. 2012.

[HN13] Longbo Huang and M.J. Neely. Utility optimal scheduling in energy-harvesting networks. *IEEE/ACM Trans. Networking*, 21(4):1117–1130, Aug. 2013.

[HQ14] R.Q. Hu and Yi Qian. An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems. *IEEE Commun. Mag.*, 52(5):94–101, May 2014.

[HRTA14] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser. Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective. *IEEE Wireless Commun.*, 21(3):118–127, Jun. 2014.

[HSDA14] Pavan Nuggehalli Zhouyue Pi Harpreet S. Dhillon, Ying Li and Jeffrey G. Andrews. Fundamentals of heterogeneous cellular networks with energy harvesting. *IEEE Trans. Wireless Commun.*, 13(5):2782–2797, May 2014.

[HUA] HUAWEI. Green energy solution. `http://www.huawei.com/uk/solutions/`

`go-greener/hw-076723.htm`.

[JSXA11] Han-Shin Jo, Young Jin Sang, Ping Xia, and J.G. Andrews. Outage probability for heterogeneous cellular networks with biased cell association. In *Proc. IEEE Globecom Conf.*, pages 1–5, Dec. 2011.

[J.W98] J.Walrand. *An Introduction to Queuening Networks.* Upper Saddle River, NJ: Prentice-Hall, 1998.

[JZLL15] S. Jangsher, Haojie Zhou, V.O.K. Li, and Ka-Cheong Leung. Joint allocation of resource blocks, power, and energy-harvesting relays in cellular networks. *IEEE J. Sel. Areas in Commun.*, 33(3):482–495, Mar. 2015.

[KBCH10] N. Ksairi, P. Bianchi, P. Ciblat, and W. Hachem. Resource allocation for downlink cellular OFDMA systems–Part I optimal allocation. *IEEE Transactions on Signal Processing*, 58(2):720–734, Feb. 2010.

[KdVYV12] Hongseok Kim, G. de Veciana, Xiangying Yang, and M. Venkatachalam. Distributed alpha-optimal user association and cell load balancing in wireless networks. *IEEE/ACM Trans. Networking*, 20(1):177–190, Feb. 2012.

[Kyo10] Kyocera. Potential performance of range expansion in macro-pico deployment(R1-104355). *3GPP TSG RAN WG1 Meeting-62*, Aug. 2010.

[LHWQ12] Qian Li, R.Q. Hu, Geng Wu, and Yi Qian. On the optimal mobile association in heterogeneous wireless relay networks. In *Proc. 2012 IEEE INFOCOM*, pages 1359–1367, Mar. 2012.

[LPGDlR+11] D. Lopez-Perez, I. Guvenc, G. De la Roche, M. Kountouris, T.Q.S. Quek, and Jie Zhang. Enhanced intercell interference coordination challenges in heterogeneous networks. *IEEE Wireless Commun.*, 18(3):22–30, Jun. 2011.

[LWZY11] Rongduo Liu, Wei Wu, Hao Zhu, and Dacheng Yang. M2M-oriented QoS categorization in cellular network. In *Proc. 2011 7th Int. Conf. on Wireless Commun. Networking and Mobile Computing (WiCOM)*, pages 1–5, Sep. 2011.

[MAAV14a] A. Mesodiakaki, F. Adelantado, L. Alonso, and C. Verikoukis. Energy-

efficient context-aware user association for outdoor small cell heterogeneous networks. In *Proc. 2014 IEEE Int. Conf. on Commun. (ICC)*, pages 1614–1619, Jun. 2014.

[MAAV14b] A. Mesodiakaki, F. Adelantado, L. Alonso, and C. Verikoukis. Energy-efficient user association in cognitive heterogeneous networks. *IEEE Commun. Mag.*, 52(7):22–29, Jul. 2014.

[MBS+10] R. Madan, J. Borran, Ashwin Sampath, N. Bhushan, A. Khandekar, and Tingfang Ji. Cell association and interference coordination in heterogeneous LTE-A cellular networks. *IEEE J. Sel. Areas in Commun.*, 28(9):1479–1489, Dec. 2010.

[MDO09] W. Mohr M. Dottling and A. Osseiran. *Radio Technologies and Concepts for IMT-Advanced*. Wiley, Dec. 2009.

[MHY+12] Jie Miao, Zheng Hu, Kun Yang, Canru Wang, and Hui Tian. Joint power and bandwidth allocation algorithm with QoS support in heterogeneous wireless networks. *IEEE Commun. Lett.*, 16(4), Apr. 2012.

[NLS13] D.W.K. Ng, E.S. Lo, and R. Schober. Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station. *IEEE Trans. Wireless Commun.*, 12(7):3412–3427, Jul. 2013.

[NLW12] D. Niyato, Xiao Lu, and Ping Wang. Adaptive power management for wireless base stations in a smart grid environment. *IEEE Wireless Commun.*, 19(6):44–51, Dec. 2012.

[NNB+13] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, Tang Hai, Shen Xiaodong, Yang Ning, and Li Nan. Trends in small cell enhancements in LTE advanced. *IEEE Commun. Mag.*, 51(2):98–105, Feb. 2013.

[Nok14] Nokia. 5G use cases and requirements. 2014.

[NRE] NREL. Pvwatts. `Available:http://rredc.nrel.gov/solar/calculators/pvwatts/version1/`.

[NSMV14] N. Namvar, W. Saad, B. Maham, and S. Valentin. A context-aware matching game for user association in wireless small cell networks. In *Proc. 2014 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*

(ICASSP), pages 439–443, May 2014.

[NZ12] Qiang Ni and C.C. Zarakovitis. Nash bargaining game theoretic scheduling for joint channel and power allocation in cognitive radio systems. *IEEE J. Selected Areas in Commun.*, 30(1):70–81, 2012.

[PBS+13] F. Pantisano, M. Bennis, W. Saad, S. Valentin, M. Debbah, and A. Zappone. Proactive user association in wireless small cell networks via collaborative filtering. In *Proc. 2013 Asilomar Conf. on Signals, Systems and Computers*, pages 1601–1605, Nov. 2013.

[PLL+11] Chunyi Peng, Suk-Bok Lee, Songwu Lu, Haiyun Luo, and Hewu Li. Traffic-driven power saving in operational 3G cellular networks. In *Proc. 17th Annual Int. Conf. Mobile Comput. and Netw.*, MobiCom '11, pages 121–132, 2011.

[PMN13] Haris Pervaiz, Leila Musavian, and Qiang Ni. Joint user association and energy-efficient resource allocation with minimum-rate constraints in two-tier HetNets. In *2013 IEEE 24th Int. Sym. on Personal Indoor and Mobile Radio Commun. (PIMRC)*, pages 1634–1639, Sep. 2013.

[PS98] C. H. Papadimitriou and K. Steiglitz. *Combinational Optimization: Algorithm and Complexity*. Dover Publicatons, 1998.

[RF14] J.B. Rao and A.O. Fapojuwo. A survey of energy efficient resource management techniques for multicell cellular networks. *IEEE Commun. Surveys Tutorials*, 16(1):154–180, First 2014.

[RLB07] B. Radunovic and J.-Y. Le Boudec. A unified framework for max-min and min-max fairness with applications. *IEEE/ACM Trans. Networking*, 15(5):1073–1083, Oct 2007.

[RPIdOV14] J. Rubio, A. Pascual-Iserte, J. del Olmo, and J. Vidal. User association for load balancing in heterogeneous networks powered with energy harvesting sources. In *Proc. 2014 Globecom Workshops (GC Wkshps)*, pages 1248–1253, Dec. 2014.

[RT10] J. Ros and W.K. Tsai. A lexicographic optimization framework to the flow control problem. *IEEE Trans. Inf. Theory*, 56(6):2875–2886, Jun.

2010.

[San] Sandvine. 2010 mobile internet phenomena report. `Available:www.sandvine.com/Content-\Pages/52357445.pdf`.

[SHL15] R. Sun, M. Hong, and Z. Luo. Joint downlink base station association and power control for max-min fairness: Computation and complexity. *IEEE J. Sel. Areas in Commun.*, PP(99):1–1, 2015.

[SHZ+14] W. Saad, Zhu Han, Rong Zheng, M. Debbah, and H.V. Poor. A college admissions game for uplink user association in wireless small cell networks. In *Proc. 2014 IEEE INFOCOM*, pages 1096–1104, Apr. 2014.

[SJL+13] M.Z. Shafiq, L. Ji, A.X. Liu, J. Pang, and J. Wang. Large-scale measurement and characterization of cellular machine-to-machine traffic. *IEEE/ACM Trans. Networking*, 21(6):1960–1973, Dec. 2013.

[SKEP12] Zukang Shen, A. Khoryaev, E. Eriksson, and Xueming Pan. Dynamic uplink-downlink configuration and interference management in TD-LTE. *IEEE Commun. Mag.*, 50(11):51–59, Nov. 2012.

[SKYK11] Kyuho Son, Hongseok Kim, Yung Yi, and B. Krishnamachari. Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. *IEEE J. Sel. A. Commun.*, 29(8):1525–1536, Sep. 2011.

[SNSD13] K. Son, S. Nagaraj, M. Sarkar, and S. Dey. QoS-aware dynamic cell reconfiguration for energy conservation in cellular networks. In *Proc. 2013 IEEE Wireless Commun. and Networking Conf. (WCNC),*, pages 2022–2027, Apr. 2013.

[SQKS13] Yong Sheng Soh, T.Q.S. Quek, M. Kountouris, and Hyundong Shin. Energy efficient heterogeneous cellular networks. *IEEE J. on Sel. Areas in Commun.*, 31(5):840–850, May 2013.

[SSV+14] O. Semiari, W. Saad, S. Valentin, M. Bennis, and B. Maham. Matching theory for priority-based cell association in the downlink of wireless small cell networks. In *Proc. 2014 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 444–448, May 2014.

[Std10] 3GPP Std. Summary of the description of candidate eICIC solutions (R1-104968). Aug. 2010.

[SY14] Kaiming Shen and Wei Yu. Distributed pricing-based user association for downlink heterogeneous cellular networks. *IEEE J. Sel. Areas in Commun.*, 32(6):1100–1113, Jun. 2014.

[SYXM13] Liyan Su, Chenyang Yang, Zhikun Xu, and A.F. Molisch. Energy-efficient downlink transmission with base station closing in small cell networks. In *Proc. 2013 IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, pages 4784–4788, May 2013.

[SZZH14] Yanzi Song, Ming Zhao, Wuyang Zhou, and Hui Han. Throughput-optimal user association in energy harvesting relay-assisted cellular networks. In *Proc. 2014 Sixth Int. Conf. on Wireless Commun. and Signal Processing (WCSP)*, pages 1–6, Oct. 2014.

[WKLY15] B. Wang, Q. Kong, W. Liu, and L.T. Yang. On efficient utilization of green energy in heterogeneous cellular networks. *IEEE Systems Journal*, PP(99):1–12, 2015.

[XHWW14] Yiran Xu, R.Q. Hu, Lili Wei, and Geng Wu. QoE-aware mobile association and resource allocation over wireless heterogeneous networks. In *Proc. 2014 IEEE Global Commun. Conf. (GLOBECOM)*, pages 4695–4701, Dec. 2014.

[YC02] Wei Yu and J.M. Cioffi. FDMA capacity of gaussian multiple-access channels with ISI. *IEEE Trans. Commun.*, 50(1):102–111, Jan. 2002.

[YMR00] H. Yaiche, R.R. Mazumdar, and C. Rosenberg. A game theoretic framework for bandwidth allocation and pricing in broadband networks. *IEEE/ACM Trans. Networking*, 8(5):667–678, Oct 2000.

[YRC+13] Qiaoyang Ye, Beiyu Rong, Yudong Chen, M. Al-Shalash, C. Caramanis, and J.G. Andrews. User association for load balancing in heterogeneous cellular networks. *IEEE Trans. Wireless Commun.*, 12(6):2706–2716, Jun. 2013.

[YZZJ11] Youwen Yi, Jin Zhang, Qian Zhang, and Tao Jiang. Spectrum leasing

to multiple cooperating secondary cellular networks. In *Proc. IEEE Int. Conf. Commun. (ICC)*, pages 1–5, Jun. 2011.

[ZHW+12] Kan Zheng, Fanglong Hu, Wenbo Wang, Wei Xiang, and M. Dohler. Radio resource allocation in LTE-advanced cellular networks with M2M communications. *IEEE Commun. Mag.*, 50(7):184–192, 2012.

[ZPSY13] Meng Zheng, P. Pawelczak, S. Stanczak, and Haibin Yu. Planning of cellular networks enhanced by energy harvesting. *IEEE Commun. Lett.*, 17(6):1092–1095, Jun. 2013.

[ZWC12] Hang Zhu, Shaowei Wang, and Dageng Chen. Energy-efficient user association for heterogenous cloud cellular networks. In *Proc. 2012 IEEE Globecom Workshops (GC Wkshps)*, pages 273–278, Dec. 2012.