December 2015

Power-Aware Planning and Design for Next Generation Wireless Networks

Chenfei Gao
Syracuse University

Abstract

Mobile network operators have witnessed a transition from being voice dominated to video/data domination, which leads to a dramatic traffic growth over the past decade. With the 4G wireless communication systems being deployed in the world most recently, the fifth generation (5G) mobile and wireless communication technologies are emerging into research fields. The fast growing data traffic volume and dramatic expansion of network infrastructures will inevitably trigger tremendous escalation of energy consumption in wireless networks, which will result in the increase of greenhouse gas emission and pose ever increasing urgency on the environmental protection and sustainable network development. Thus, energy-efficiency is one of the most important rules that 5G network planning and design should follow.

This dissertation presents power-aware planning and design for next generation wireless networks. We study network planning and design problems in both offline planning and online resource allocation. We propose approximation algorithms and effective heuristics for various network design scenarios, with different wireless network setups and different power saving optimization objectives. We aim to save power consumption on both base stations (BSs) and user equipments (UEs) by leveraging wireless relay placement, small cell deployment, device-to-device communications and base station consolidation.

We first study a joint signal-aware relay station placement and power allocation problem with consideration for multiple related physical constraints such as channel capacity, signal to noise ratio requirement of subscribers, relay power and

network topology in multihop wireless relay networks. We present approximation schemes which first find a minimum number of relay stations, using maximum transmit power, to cover all the subscribers meeting each SNR requirement, and then ensure communications between any subscriber and a base station by adjusting the transmit power of each relay station. In order to save power on BS, we propose a practical solution and offer a new perspective on implementing green wireless networks by embracing small cell networks. Many existing works have proposed to schedule base station into sleep to save energy. However, in reality, it is very difficult to shut down and reboot BSs frequently due to numerous technical issues and performance requirements. Instead of putting BSs into sleep, we tactically reduce the coverage of each base station, and strategically place microcells to offload the traffic transmitted to/from BSs to save total power consumption.

In online resource allocation, we aim to save tranmit power of UEs by enabling device-to-device (D2D) communications in OFDMA-based wireless networks. Most existing works on D2D communications either targeted CDMA-based single-channel networks or aimed at maximizing network throughput. We formally define an optimization problem based on a practical link data rate model, whose objective is to minimize total power consumption while meeting user data rate requirements. We propose to solve it using a joint optimization approach by presenting two effective and efficient algorithms, which both jointly determine mode selection, channel allocation and power assignment.

In the last part of this dissertation, we propose to leverage load migration and base station consolidation for green communications and consider a power-efficient network planning problem in virtualized cognitive radio networks with the objective of minimizing total power consumption while meeting traffic load demand of each Mobile Virtual Network Operator (MVNO). First we present a

Mixed Integer Linear Programming (MILP) to provide optimal solutions. Then we present a general optimization framework to guide algorithm design, which solves two subproblems, channel assignment and load allocation, in sequence. In addition, we present an effective heuristic algorithm that jointly solves the two subproblems.

Numerical results are presented to confirm the theoretical analysis of our schemes, and to show strong performances of our solutions, compared to several baseline methods.

Power-Aware Planning and Design for Next Generation Wireless Networks

by

Chenfei Gao

B.S., Beijing University of Posts and Telecommunications (2009)
M.S., Syracuse University(2011)

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor Of Philosophy in Electrical And Computer Engineering

Syracuse University
December 2015

Acknowledgements

I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family and wife.

First, and foremost, I would like to express my deepest gratitude to my advisor, Dr. Jian Tang and my mentor Dr. Weiyi Zhang, for their excellent guidance, caring, patience, motivation, and providing me with an excellent atmosphere for doing research. Their guidance helped me a lot across all the time of doing research and writing this thesis. I could not have imagined to have a better advisor and mentor for my Ph.D study.

I would like to thank all my labmates: Jielong Xu, Jimmy Chen, Jing Wang, Teng Li and Zhiyuan Xu. They are not only labmates but also my close friends. Many thanks for all your collaboration, company and encouragement. The past years, without you, would be a lonely research journey. You made the time much more enjoyable.

I would also like to acknowledge the financial support from NSF grants CNS-1113398 and CNS-1443966.

My final but everlasting gratitude goes to my family members. They were always supporting me and encouraging me with their best wishes. Special thanks goes to my wife, Yufei Wu. She was always there cheering me up and stood by my through the good times and bad.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Next Generation Wireless Networks

1.1.1 Vision and Prospects

Mobile wireless communications have experienced explosive growth over the past decade, fueled by the popularity of smartphones, tablets and diverse cloud services. Mobile network operators have witnessed a transition from being voice dominated to video/data domination. A broad consensus in the wireless industry anticipates a strong contribution of this trend for several years to come. With the maturing of fourth generation (4G) standardization and the ongoing worldwide deployment of 4G cellular networks, technologists worldwide have begun searching for next generation wireless solutions to meet the anticipated demands in the 2020 era given the explosive growth of mobile Internet. New research projects such as METIS, iJOIN and 5GNOW, have started internationally, and research centers such as Samsung, the Fraunhofer Heinrich Hertz Institute and the European Telecommunications Standards Institute, devoted to 5G technology have begun to open [77].

The development of 3G and 4G wireless technology was mainly driven by demand for data services over the Internet. However, the drivers for 5G systems are likely to be much more diverse. It is widely agreed that compared to the 4G network, the 5G network should achieve 1000 times the system capacity, 10 times the spectral efficiency, energy efficiency and data rate (i.e., peak data rate of 1 Gb/s for low mobility), and 25 times the average cell throughput [80]. The 5G network aims to connect the entire world, and achieve seamless and ubuquitous communications between anybody, anything, wherever they are, whenever they need, by whatever electronic devices/services/networks they wish.

1.1.2 Key Technologies

In this section, we discuss some promising key wireless technologies that can enable 5G wireless networks to fulfill performance requirements. The purpose of developing these technologies is to build an energy efficient 5G network with a dramatic capacity increase and efficient utilization of all possible resources.

- **Energy Efficiency (EE) and Spectral Efficiency (SE) co-design**

  Given limited spectrum and ever increasing capacity demand, SE has been pursued for decades as the top design priority of all major wireless standards, ranging from cellular networks to local and personal area networks. SE-oriented designs, however, have overlooked the issues of infrastructure power consumption. Currently, RANs consume 70 percent of the total power [42]. A sustainable future wireless network must therefore be not only spectral efficient but also energy efficient. Hence, EE and SE joint optimization is a critical part of 5G research.

- **Spatial densification via small cells and relays**

  To handle the explosively growth of traffic demands, network operators need to deploy additional macro base stations (BSs), leading to significant cost and elaborate site planning. Low-power nodes (i.e. small cells which may be employed indoors or outdoors) offer a simpler cost-effective alternative to conventional cell splitting. Small form factor and low power rating of small cells (e.g. microcells, picocells) enable much lower capital expendiure (CAPEX) and operational expenditure (OPEX) compared to macrocells. At locations without wired backhaul access, relay nodes may be deployed instead of small cells. A relay node uses wireless/cellular spectrum not only to provide access to mobile users, but also to backhaul data to an anchor BS with wired backhaul [12].

- **Device-to-Device (D2D) communications**

  In voice-centric systems it was implicitly accepted that two parties willing to establish a call would not be in close proximity. In the age of data, this premise might no longer hold, and it could be common to have situations where several co-located devices would like to wirelessly share content or interact [14]. D2D communication commonly refers to the technique that enables wireless devices to communicate directly with each other without an infrastructure of access points or BSs. The BS only helps user devices set up connections without relaying any data traffic. D2D communication can potentially improve user experience by reducing latency and power consumption, increasing peak data rates, and creating new proximity-based services such as proximate multiplayer gaming [12]. Thus, it has been considered as a key enabling technology for the next generation (i.e. 5G) wireless communications.

- **Cognitive radio (CR) technology**

  Emerging CR technology and the Dynamic Spectrum Access (DSA) apprach [4] enable unlicensed wireless users (a.k.a secondary users) to sense and access the under-utilized spectrum opportunistically even if it is licensed. The CR networks should be aware of the surrounding radio environment and regulate its transmission accordingly. Adopting CR is motivated by the fact that a large portion of the radio spectrum is underutilized most of the time [80]. It has been considered as the next generation wireless communication technology that can lead to better spectrum utilization and higher network capacity.

1.2 Energy-Efficient Network Planning and Design

1.2.1 Motivation

As described in Section 1.1, energy-efficiency is one of the most important aspects in planning and design of next generation wireless networks. Due to rapid growth of wireless terminals and their traffic demands, wireless networks have become one of the largest contributors to power usages. Information and communications technologies (ICT) takes up a considerable proportion of total energy consumption. In 2012, the annual average power consumption by ICT industries was over 200 GW, of which telecoms infrastructure and devices accounted for 25 percent [74]. ICT already represents around 2% of total carbon emissions, of which mobile networks represent about 0.2%, and this is expected to increase every year. In the 5G era, it is expected that millions more base stations (BSs) with higher functionality and billions more smart phones and devices with much higher data rates will be connected. The largest mobile network in the world consumed over 14 billion kWh of energy in 2012 in its network of 1.1 million BSs [42]. Such huge energy consumption has raised public concerns about electricity costs, and greenhouse gas emissions that are known to have a negative impact on global climate.

As global carbon emissions increase and sea levels rise, global weather and air pollution in many large cities across the world is becoming more severe [34]. In addition to the environmental aspects, energy costs also represent a significant portion of network operators overall expenditures (OPEX). The rising energy costs and carbon footprint of operating cellular networks have led to an emerging trend of addressing energy-efficiency amongst the network operators and regulatory bodies such as 3GPP and ITU [1, 45]. Therefore, the problem of how to build a green (power-efficient) wireless network has attracted extensive research attention from both industry and academia recently. There is significant potential for power savings in wireless networks. European Commission has recently started new projects within its seventh Framework Programme to address the energy efficiency of mobile communication systems, viz. energy Aware Radio and NeTwork TecHnologies (EARTH), Towards Real Energy-efficient Network Design (TREND), and Cognitive Radio and Cooperative strategies for Power saving in multi-standard wireless devices (C2POWER) [15, 24, 78]. If green communications technologies are universally deployed across current/future network, significant energy savings can be realized, enabling larger infrastructure deployments for 4G and 5G capacity upgrades without requiring significant increase in average revenue per user (ARPU).

Most previous research efforts, however, were mainly focused on reducing energy consumption of battery powered wireless devices such as mobile phones and sensor nodes. Research attention has not been well paid to power savings on BSs until very recently. The most straightforward way to reduce power consumption of BSs is to turn off idle BSs or put them into sleep. However, without careful network planning, turning off BSs might lead to loss of coverage and unsatisfied traffic demands.

Considering power saving has been recognized as an urgent issue worldwide, we aim to conduct some research works focusing on saving power consumption on both wireless infrastructure (e.g. macro-BS, micro-BS, relay nodes) and user equipments (UEs). By embracing some of key promising technologies in next generation wireless networks such as small cell network deployment and D2D communications, we study the optimization problems in both offline network planning to save the power consumption on macro-BSs, micro-BSs, relay nodes and online resource allocation to save the aggregate transmit power consumption on UEs in potential D2D links.

It is known that small cell network is one of these new trends for next generation mobile network design. One model is using Relay Stations (RS) as small cell providers to achieve extended coverage, lower cost, and higher network capacity. In offline network planning, we first study a joint signal-aware RS placement and power allocation problem across a field of multiple BSs in wireless relay networks. We aim to design a low-cost multihop wireless relay network with the consideration of channel capacity, subscriber's SINR requirement, power consumption of relay nodes and multi-cell scenario. However, saving the power consumption on RSs is not enough. In order to save more power, a practical solution needs to be brought forward to save the power consumption on macro-cell BSs. Unlike previous works in literature in which idle BSs or under-utilized BSs are turned off or put into sleep to save power, we aim to propose a more practical solution and offer a new perspective on implementing green wireless networks by embracing the hot trending small cell networks. We would like to see how (or if) small cell networks can provide a more energy-efficient greener wireless network.

Besides the research in offline planning and design, we also intent to study the online resource allocation problem in the context of D2D underlaying cellular networks as considering D2D communications are very likely to take place in next generation wireless networks. BSs need to determine the transmission mode for each potential D2D link and also allocate resources for them effectively and efficiently. Thus, the joint optimization problem of mode selection and resource allocation is worth studying. Our goal is to save the total transmit power of UEs in potential D2D links such that UEs can have longer battery lifetime.

At last, we combine wireless resource virtualization with green wireless network and then claim that multiple Mobile Virtual Network Operators (MVNOs) can be supported over a shared physical wireless network and traffic loads in a BS can be easily migrated to more power-efficient BSs in its neighborhood such that idle BSs can be turned off or put into sleep to save power. We propose to leverage load migration and BS consolidation for green communications and consider a power-efficient network planning problem in virtualized Cognitive Radio Networks (CRNs) with the objective of minimizing total power consumption while meeting traffic load demand of each MVNO.

All our research works are targeting the energy-efficiency in planning and design of next generation wireless networks. Our ultimate goal is to contribute novel ideas and effective solutions to the research community of green wireless networking.

1.2.2 State of The Art and Literature Gap

In this section, we describe the state of the art and discuss the literature gap on the areas that we work on, which are listed as below.

• **Power efficiency on wireless infrastructure**

Green communications and networking, especially power efficiency on BSs and wireless infrastructure, has attracted many researchers' attentions recently. [39] and [16] are two surveys summarizing the current research works on green wireless network, especially on green cellular networks. Many BS equipment manufacturers have begun to offer a number of eco and cost friendly solutions to reduce power demands of BSs and to support off-grid BSs with renewable energy resources. Nokia Siemens Networks Flexi Multiradio Base Station, Huawei Green Base Station and Flexenclosure Esite solutions are examples of such recent efforts [28, 41, 66].

In current literatures, the most straightforward way to reduce power consumption of BSs is to turn off idle BSs or put them into sleep. [23, 25, 30, 60, 61, 69, 72] are all working on shutting down some of under-utilized BSs in order to achieve the power savings via different approaches while satisfying various constraints in the network. In [69], Peng et al. proposed a profile-based approach to green cellular infrastructure, which leveraged emporal-spatial traffic diversity and node deployment heterogeneity, and powered off under-utilized BSs based on historical data. The authors of [30] provided an algorithm that minimized power consumption by selectively turning on or off cell towers and deciding which power to assign to the active nodes and what frequencies to use, so as to maintain full coverage and respect users capacity demands. In [25], Elayoubi et al. investigated network sleep mode for reducing energy consumption of radio access networks. They proposed an offline-optimized controller that associated traffic with an activation/deactivation policy that maximized a multiple objective function of QoS and energy consumption. In [60], the authors showed how to optimize the energy saving, first assuming that any fraction of cells could be turned off, and then accounting for the constraints resulting from the cell layout. [23] proposed a concrete methodology for saving energy, which was based on re-arranging the user-cell association so as to allow shutting down under-utilized parts of the network. [61] investigated the energy saving potential of exploiting cell size breathing by putting low loaded cells in to sleep mode. In [72], the authors developed a system selection algorithm that found the optimal traffic allocation for the different systems that minimized power consumption while insuring the target QoS. Besides utilizing sleep mode to save energy, the authors in [10] studied the effect of cell sizes on the energy consumption and proposed a practical, 2-level scheme that adjusted cell sizes between two fixed values, and showed an energy saving of up to 40%. In [73], the authors first studied how to adaptively vary the processing speed based on incoming load and then proposed and analyzed a distributed algorithm, called Speed Balance, that could yield significant energy savings. However, without careful network planning, turning off BSs might lead to loss of coverage and unsatisfied traffic demands. Furthermore, in reality, it is very difficult to shut down and reboot BSs frequently due to numerous technical issues and performance guarantees. Shutting down BSs would lead to loss of coverage due to handover delays, which were not carefully considered in above mentioned works. Rebooting BSs requires air conditioners to spend extra power on heating to indoor temperature. The extra power consumption was usually neglected in most works. Hence, in practice, to find a feasible short-term solution, we need to seek other opportunities.

- **Relay station placement**

Small cell network is one of many new trends for next generation wireless networks since many mobile network operators see small cells as vital to managing spectrum more efficiently. Ideally, small cell network scheme can help network carriers to achieve extended coverage and higher network capacity. One of the feasible small cell network designs is using RS to offload traffic that directly transmitted to/from macro cells.

Relay station placement has been an active research topic in wireless networks, especially in wireless sensor networks. By using RSs, one could deploy a network at a lower cost than using only (more expensive) BSs to provide wide coverage while delivering a required level of service to users [33, 54, 55, 76]. In [56], Lin and Xue proved the single-tiered placement problem with $R = r$ and $K = 1$ was NP-hard, where R, r and K denoted the transmission range of relay nodes, the transmission range of sensor nodes, the connectivity requirement respectively. A 5-approximation algorithm was presented to solve the problem. The authors also designed a steinerization scheme which had been used by many later works. Beside minimizing the number of placed RSs, extensive research has been done on placement with physical constraints, such as energy consumption and network lifetime. Hou et al. studied the energy provisioning problem for a two-tiered wireless sensor network [40]. Besides provisioning additional energy on the existing nodes, they considered deploying relay nodes (RNs) into the network to mitigate network geometric deficiency and prolong network lifetime. In [84], Hassanein et al. proposed three random relay deployment strategies for connectivity-oriented, lifetime-oriented and hybrid deployment. In [67], Pan et al. studied BS placement to maximize network lifetime. Recently, a new dual-relay coverage architecture was proposed for 802.16j Mobile Multi-hop Relay-based (MMR) networks [54,55], where each subscriber station (SS) was covered by two RSs. [54] assumed that only one RS was placed in each cell. ILP formulation was applied to find an optimal placement of RS which could maximize the cell capacity in terms of user traffic rates. In [55], assuming a uniform distribution on user traffic demand, the authors studied how to determine the RSs locations from a set of predefined candidate positions. Quality of service provisioning in telecommunications networks has been shown to be important to study in practice [13]. Considering channel quality, the authors of [90] studied multiple hop relay problem. Two tiers model was mentioned as well, but it addressed the relay placement problem on condition that all relay nodes forwarded traffic in their maximum transmit power. In addition, an efficient MUST algorithm was proposed to address the connectivity problem on upper tier. However, MUST worked under the physical constraint of only one BS in the field. Thus by far, limited work has studied relay station placement problem in a more general wireless relay networks of multiple BSs, considering a practical set of physical constraints such as indivisual channel capacity, SNR requirement of subscriber, relay power and network topology together.

• **Mode selection and resource allocation in D2D communications**

In the context of D2D communications, resource allocation has been addressed by quite a few research works recently. In [3], the authors proposed a coalitional game based approach for mode selection of D2D links, with the objective of minimizing the total power while satisfying rate requirements. In [6], the authors targeted at the energy-efficient resource allocation for D2D communications as an underlay of a fully loaded LTE-A network. They aimed to maximize the number of admissible D2D pairs thereby minimizing the total uplink transmit power of cellular and D2D links by solving two subproblems separately. The authors of [17] formulated a joint optimization problem as a Mixed Integer Non-Linear Programming (MINLP) problem, where the mode to operate, radio resources to use, and power to transmit were to be optimally decided for a group of users. They also presented a heuristic algorithm with reduced complexity. The authors of [22] proposed novel mode selection algorithms that took into account the interference situation and the operational state of the cellular network in both single-cell and multi-cell scenarios. In [37], the authors derived means for obtaining optimal communication modes for all devices in the system in terms of system equations, which captured network states such as link gains, noise levels, SINR, etc. Furthermore, practical communication mode selection algorithms were presented to show their performance against the achievable bounds. In [38], Han et al., developed a stochastic framework for sub-channel and transmission mode scheduling, with the objective of maximizing the average sum-rate of the system, while satisfying the Quality-of-Service (QoS) requirement of each user. In [46], the authors proposed an exhaustive search based mode selection and power assignment scheme for D2D communication systems. The proposed scheme searched all possible mode combinations which consisted of mode indices for all devices in the system. In [53], the authors formulated the joint mode selection and resource allocation in D2D communications underlaying cellular networks as a flow maximization problem based on the transmission graph and then optimally solved it. In [57], the underlay and overlay mode selections were analyzed for D2D communications in the LTE-advanced single-cell scenario. Their results showed that the underlay mode was preferred when the cellular user was closer to the BS or relay node than the D2D user. In [82, 83], Xiang et al. , presented a distance-dependent algorithm and a cooperative mode selection mechanism respectively, both aiming at selecting optimal transmission modes with overall capacity maximized and QoS of mobile users satisfied. In [85], Yu et al. analyzed optimum resource allocation and power control, aiming to optimize throughput over shared resources while fulfilling prioritized cellular service constraints. It was found that in most of the cases, optimum power control and resource allocation for resource sharing modes could either be solved in closed form or searched from a finite set. In [86], the authors considered joint mode selection, channel assignment, and power control to maximize overall system throughput. They decomposed the optimization problem into two subproblems: power control, joint mode selection and channel assignment. They developed low-complexity heuristic algorithms to solve the subproblems. The authors of [92] proposed a dynamic stackelberg game framework in which the BS and potential D2D UEs acted as the leader and the followers respectively to jointly address the problems of spectrum allocation and user-controlled mode selection. The authors of [91] studied the joint optimization problem of D2D mode selection, modulation and coding scheme assignment, resource block and power allocation with the objective of minimizing the overall power consumption under rate requirements. They decoupled the problem into two sub-problems, which were solved by Lagrangian relaxation and tabu search methods, respectively. General introductions to D2D communications, and related protocols and standards can be found in [21, 51]. However, limited work has addressed joint resource optimization problems in the context of D2D communications and OFDMA-based wireless networks based on a practical link data rate model with the objective to minimize total power consumption.

- **Wireless resource virtualization**

Even though server/desktop virtualization has been well studied, research on wireless resource virtualization is still in its infancy. In [48], Kokku et al. described the design and implementation of a Network Virtualization Substrate (NVS) for effective virtualization of wireless resources in cellular networks. NVS met three key requirements: isolation, customization, and efficient resource utilization. They demonstrated its efficacy via a prototype implementation and evaluation on a WiMAX testbed. In [65], the authors proposed a Cognitive Virtualization Platform called AMPHIBIA, which enabled end-to-end slicing over wired and wireless networks and exploited the network advantages of virtualization and CR technologies. In [93], Zhu et al. introduced the first TDD WiMAX Software Defined Radio (SDR) BS implemented on a commodity server, in conjunction with a novel design of a remote radio head. In [52], the authors presented a software-defined cellular network architecture that supported flexible slicing of network resources. Wireless resource virtualization has also been studied for LTE networks [89], WiMAX networks [9], WiFi networks [81], access networks [49], multihop wireless networks [88], and wireless sensor networks [44]. However, most of works were mainly focused on how to design and implement resource virtualization at one node. Limited work has studied BS consolidation (that can be enabled by virtualization) for reducing power consumption of the whole network in the context of a network with virtualized cognitive radio BSs that are shared by multiple MVNOs.

1.2.3 Contributions

In this section, we present our contributions from the research works conducted in this thesis. Overally speaking, we provide effective solutions to the offline network planning and design problems by leveraging wireless relay placement and small cell deployment in the context of wireless relay networks and small cell networks, respectively. Additionally, we enable green D2D communications in OFDMA-based wireless networks via a joint optimization approach. We effectively and efficiently solve the online joint optimization problem of mode selection, channel allocation and power assignment for potential D2D links, and also demonstrate the substantial power savings by our approach, compared to several baseline methods. At last, in virtualized cognitive radio networks, we effectively solve a power-efficient network planning problem by leveraging load migration and BS consolidaion for green communications.

More detailedly, our contributions arise from three major research works. In offline network planning, we provide effective solutions to base/relay station placement in heterogenous cellular networks. In online resource allocation, we study joint resource optimization for green D2D communications and explore BS consolidation in virtualized cognitive radio networks, respectively.

• **Base/relay station placement in heterogenous cellular networks**

For base/relay station placement in heterogenous cellular networks, we first study a joint signal-aware RS placement and power allocation problem with multiple BSs in wireless relay networks with taking into account multiple related physical constraints such as channel capacity, signal to noise ratio (SNR) requirement of subscribers, relay power and network topology. We present approximation schemes which first find a minimum number of RS, using maximum transmit power, to cover all the subscribers meeting each SNR requirement, and then ensure communications between any subscriber to a BS by adjusting the transmit power of each RS. Numerical results are presented to confirm the theoretical analysis of our schemes, and to show strong performances of our solutions. Then, to save power consumption on macro-BS, we propose a much more practical solution and offer a new perspective on implementing green wireless networking by embracing the hot-trending small cell network idea. Instead of putting BSs into sleep, we tactically reduce the coverage (and the power usage) of each BS, and strategically place microcells (relay stations) to offload the traffic transmitted to/from BSs. We propose approximation algorithms for various network design scenarios, with different wireless network setups and different power saving optimization objectives. Extensive numerical results have been conducted to support our theoretical analysis and showed that our schemes can provide up to 52% network power consumption compared to traditional wireless macro cell networks.

• **Joint resource optimization for green D2D communications**

For online optimization for green D2D communications, we formally define an optimization problem for power-efficient D2D communications in OFDMA-based wireless networks based on a practical link data rate model. We present two effective and efficient algorithms to solve it in polynomial time, which both jointly determine mode selection, power assignment and channel allocation. It has also been shown by extensive simulation results that the proposed algorithms can achieve over 68% power savings, compared to several baseline methods.

• **Base station consolidation in virtualized cognitive radio networks**

In BS consolidation work, we propose to leverage load migration and BS consolidation for green communications and consider a power-efficient network planning problem in virtualized cognitive radio networks with the objective of minimizing total power consumption while meeting traffic load demand of each MVNO. We formally define a BS consolidation problem and present an Mixed Integer Linear Programming (MILP) formulation to provide optimal solutions. We present a general optimization framework to guide algorithm design, which solves two subproblems, channel assignment and load allocation, in sequence. We present a channel assignment algorithm with an approximation ratio of $(\frac{1}{\Delta})$ (where Δ is the maximum number of BSs a BS can potentially interfere with). For the load allocation problem, we present a polynomial-time optimal algorithm for a special case where BSs are power-proportional as well as two fast heuristic algorithms for the general case. It has been shown by extensive simulation results that the proposed algorithms produce close-to-optimal solutions, and moreover, achieve over 45% power savings compared to a baseline algorithm that does not migrate loads or consolidate BSs.

1.3 Outline Of This Thesis

The rest of this thesis is organized as follows:

We present our work on base/relay station placement in heterogenous cellular networks in Chapter 2. In this Chapter, we first focus on relay station placement problems and aim to design a multihop green wireless relay network with the consideration of some practical physical constraints, which is presented in Section 2.1. Then we present our work on microcell BS placement in Section 2.2. We aim to reduce the coverage of macrocell BSs and meanwhile deploy microcell BSs to offload the traffic transmitted to/from macrocell BSs in order to save total power consumption.

Besides saving power consumption on wireless infrastructure by deploying small cells, we also conduct research on saving total transmit power consumption on UEs. We consider a green wireless network with D2D links and study a joint optimization problem of mode selection and resource allocation in D2D underlaying cellular networks, which is presented in Chapter 3.

In Chapter 4, we present our work to save power consumption of BSs in virtualized cognitive radio networks by leveraging load migration and BS consolidation. In such networks, multiple mobile virtual network operators can be supported over a shared physical wireless infrastructure and traffic loads in a BS can be easily migrated to more power-efficient BSs in its neighborhood such that idle BSs can be turned off or put into sleep to save power.

The conclusions of this thesis and future works are presented in Chapter 5.

Chapter 2

Base/Relay Station Placement in Heterogenous Cellular Networks

In this chapter, we introduce and discuss our research work on the topic of base/relay station placement in heterogenous wireless networks. It consists of two parts: relay station placement and microcell BS placement. First, we design a low-cost multihop wireless relay network with the consideration of practical physical constraints such as channel capacity, subscriber's SNR requirements, power consumption of relay nodes and multi-cell scenario. In the relay network, relay stations are placed to provide service to subscribers and ensure communications between any subscriber to a BS by adjusting the transmit power of each relay station. Then we bring forward a new idea and propose a practical solution to save the power consumption on macrocell BSs without turning off idle or under/utilized BSs. We claim to place small cells (e.g. microcells) strategically and tactically reduce the coverage (and the power usage) of each macrocell BS with the objective of minimizing the total power consumption.

The rest of this chapter is organized as follows. We discuss SNR-aware relay station placement in Section 2.1 including network model in Section 2.1.2, problem statements in Section 2.1.3, approximation algorithms from Section 2.1.4 to Section 2.1.6 and numerical results in Section 2.1.7. The work on power-aware microcell BS placement is introduced in Section 2.2. The corresponding problem statements, solutions and numerical results are presented from Section 2.2.2 to Section 2.2.3.

2.1 SNR-Aware Relay Station Placement

2.1.1 Overview

With the exponential growth in mobile data traffic, how to better utilize the spectrum and improve network throughput has been an important issue in telecommunication. Many are using WiFi data offloading as a more efficient use of radio spectrum. Others are looking into how to improve network capacity by better reusing spectrums. Small cell network is one of many new trends for next generation wireless networks since many mobile network operators see small cells as vital to managing spectrum more efficiently. Ideally, small cell network scheme can help network carriers to achieve extended coverage and higher network capacity. One of the feasible small cell network designs is using Relay Stations (RS) for offloadindg traffic that directly transmitted to/from macro cells to achieve extended coverage, lower cost, and higher network capacity.

In this work, we study a joint signal-aware RS placement and power allocation problem with multiple BSs in wireless relay networks considering multiple related physical constraints such as channel capacity, signal to noise ratio (SNR) requirement of subscribers, relay power and network topology. we extended the research of [32] by considering different *SNR* thresholds to users. Each user has its own *SNR* threshold value based on its data rate request. Generally, users have higher *SNR* thresholds when higher data rate are requested. However, the *SNR* threshold considered in [32] was a constant in a range of $[-25dB, -10dB]$ for all users. This universal setting of *SNR* threshold is not practical enough since user requests are normally heterogeneous. To the best of our knowledge, this work is the first one to study low-cost multi-hop relay problem considering channel capacity, subscriber's *SNR* requirement, power consumption of relay nodes and multiple BSs in wireless multi-hop networks.

2.1.2 Network model

In our model, a wireless multi-hop network consists of *Subscriber Stations* (SS), *Base Stations* (BS), and *Relay Stations*. In reality, several types of SS exist, including static SS, adhoc SS and compound SS. In recent study, [79] has demonstrated that traffic from mobile access is less than 20%, while majority of wireless traffic is actually coming from infotainment (such as video streaming, online gaming), which would not be used by mobile users regularly. It is also shown that most mobile users usually only check emails and browse web, which only contributes a small proportion of total traffic. In [27] it shows that web browsing accounts for 10% and less than 10% in 2013 and 2019, respectively. Given this character of wireless traffic, in this work, we assume that SSs are *static users* such as Wal-mart, McDonald's, and gas stations, which are static but have large traffic demands. Each SS represents the aggregated traffic coming from these service locations.

Similarly, all the RSs, with the function of relaying traffic coming from BS, other RSs, or SS, are assumed to be fixed as well in this work. Our network model divides the network into two tiers, *lower tier* and *upper tier*. In the lower tier, *coverage RSs* are placed in order to *cover* all the SSs while *meeting SS's performance requirements* such as channel capacity, *SNR* threshold. Communications in the lower tier are mainly between SSs and coverage RSs, which are denoted as "*access links*". In the upper tier, *connectivity RSs* are to be placed in order to *connect coverage RSs to BSs, using possible multiple-hop relay model*. The communication links in the upper tier are denoted as "*relay links*" in this work. The scenario described above is illustrated in Fig.2-1.

2.1.2.1 *SNR*-Aware Green Relay Allocation

Each SS needs to be covered by an RS or BS for traffic transmission. Different from most previous work, we take channel capacity and *SNR threshold* into consideration in this work.

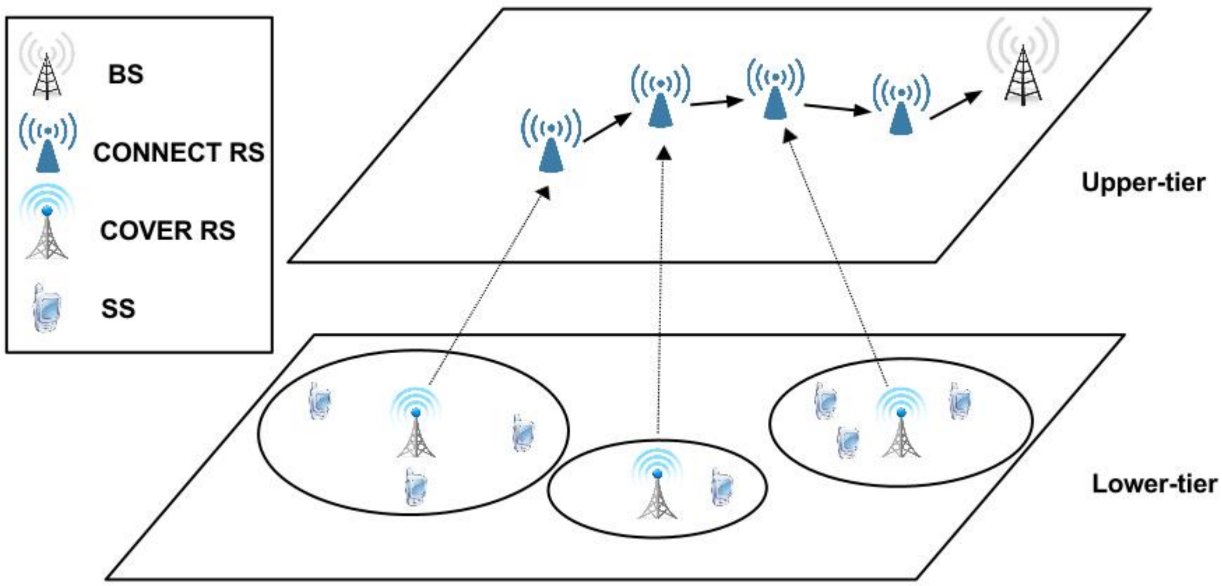


Figure 2-1: Scenario Illustration

The *access links* between an SS and its coverage RS should provide enough channel capacity to satisfy the SS's data rate request. In addition, for each SS being able to correctly decode signals, *its received* **Signal to Noise ratio (***SNR***)** *is another parameter that should be considered*. Typical 802.16 adaptive modulation and coding parameters are used to estimate the throughput achievable as a function of *SNR*. The relationship among adaptive modulation, minimum *SNR* and user throughput is listed in Table.2.1. From Table.2.1 we can see, each user needs to satisfy a minimum *SNR* threshold if its throughput reaches in the range $[10Mb/s, 45Mb/s]$. For instance, if an user has $25Mb/s$ throughput, then its received *SNR* needs to be at least $14.5dB$ so that it can correctly decode the signals. Hence, there are different *SNR* threshold values for the users with various data rate requests.

Modulation	Minimum SNR, dB	User throughput, Mb/s
QPSK 1/2	10	10
16QAM 1/2	14.5	20
16QAM 3/4	17.25	30
64QAM 2/3	21.75	40
64QAM 3/4	23	45

Table 2.1: Minimum *SNRs* with various throughputs

Definition 1 (Feasible coverage)**.** Let s_i be a fixed SS with known location, and b_i be its data rate request (in terms of *bps*). An RS r_m is said to provide *a feasible coverage* for s_i if the channel capacity of the link (in terms of *bps*) between s_i and r_m is sufficient for the data rate request of s_i; **and**, the *SNR* received at s_i is above the *SNR* threshold. □

Definition 2 (*SNR* for subscribers)**.** Let s_i be an SS, $R = \{r_1, r_2, ..., r_n\}$ be the RS set and $P = \{p_1, p_2, ..., p_n\}$ be the set of received power by s_i from each RS. If SS s_i receives signal from RS r_j, the *SNR* at s_i is defined as $\frac{p_j}{\sum_{i=1}^{n} p_i - p_j}$. □

To simplify the study, we transform the *capacity* and *SNR* requirements into *distance* requirements since the capacity of a wireless link is highly related to the distance between transmitters and receivers [18]. In this work, we choose two-ray ground path loss model for modeling the long distance LOS channel with large scale signal strength. The received power P_r is given as

$$P_r = P_t G_t G_r h_t^2 h_r^2 d^{-\alpha} \tag{1.1}$$

where P_t is the transmit power, and G_t,G_r and h_t,h_r are the gains and heights of transmitter tower and receiver tower, respectively. d is the Euclidean distance between the two end nodes, α is the attenuation factor, which usually varies in a range of $2-4$. According to Shannon's theorem, wireless link capacity is given by $C = B\log(1 + SNR_r)$, where B is the channel bandwidth. Thus, if noise N_0 is a constant, the channel capacity (in terms of bps) is only related to the received signal power P_r and moreover only related to the distance between two end nodes assuming transmit power P_t of RS is constant. Therefore, the capacity requests of SS are equivalent to distance requests between SS and its corresponding RS.

2.1.3 Problem Statements

Definition 3 (*SNR* Aware Green (SAG) Relay problem)**.** Given a wireless relay network with multiple BSs and a set of SSs $S = \{s_1, s_2, ..., s_n\}$, let $SNR = \{\beta_1, \beta_2, ..., \beta_n\}$ be the feasible *SNR* thresholds for SSs, The SAG problem seeks a minimum number of RSs R and transmit power allocation strategy for R such that:

1. Providing feasible coverage for each $s_i \in S$. Specifically, each SS $s_i \in S$ has enough *SNR* and an access link with enough capacity to an RS or BS;

2. Each placed RS must provide enough capacity on relay links to transit traffic to a BS;

3. Sum of transmit powers of the placed RSs should be minimized. □

Unlike previous coverage problems, which assume that RSs always transmit in maximum power, we allow to adjust power consumption of RSs as long as the adjustment does not change the coverage topology. A similar problem, DARP, has been studied in [90] without considering power minimization. Since DARP is estimated to be *NP*-hard [90], we expected SAG to be *NP*-hard as well. To solve SAG, our solution consists of two aspects, *coverage with minimum number of RSs*, and *minimizing transmit power of the placed RSs*. First, we assume that all the RSs are operating with maximum transmit power. With this assumption, we aim to find *a minimum number* of RSs to provide *feasible coverage* for all the SSs. In the second step, power optimization scheme will be applied to reduce the power consumptions. Naturally, we divide the original problem into two sub-problems, *Lower-tier Coverage Relay Allocation* (LCRA) problem and *Upper-tier Connectivity Relay Allocation* (UCRA) problem, which are defined in following, and try to tackle them one by one.

Definition 4 (Lower-tier Coverage Relay Allocation (LCRA) problem)**.** Given a wireless relay network with a set of subscriber station $S = \{s_1, s_2, ..., s_n\}$, and the *SNR* threshold set for SSs $SNR = \{\beta_1, \beta_2, ..., \beta_n\}$. The LCRA problem seeks *K, minimum number* of relay stations to provide *feasible coverage* for $s_i \in S$, and *the total transmit power by deploying K RSs is minimized.* □

On the upper tier, we need to consider how to transmit all the traffic from *coverage* RSs to BSs. We name the RSs placed on the upper tier *connectivity RSs* since the function of RSs in UCRA is to relay the communications between coverage RSs and BS. Similar to LCRA problem, we first assume that RSs relay with *maximum power* so that we can determine the minimum number of RS locations.

Definition 5 (Upper-tier Connectivity Relay Allocation (UCRA) problem)**.** Given a wireless relay network with a set of coverage RSs $R_c = \{r_1, r_2, ..., r_n\}$, distance requirements $D_r = \{d_r^1, d_r^2, ..., d_r^n\}$ for R_c, set of BSs $B = \{bs_1, bs_2, ..., bs_m\}$, UCRA seeks a *minimum number of connectivity RSs* operating with minimum power that ensures the communications between *coverage RSs* and BSs. □

In next sections, we will first tackle LCRA and UCRA problems separately, and then provide a solution to SAG by using a combination of the solutions to LCRA and UCRA.

2.1.4 Approximation Solutions for LCRA

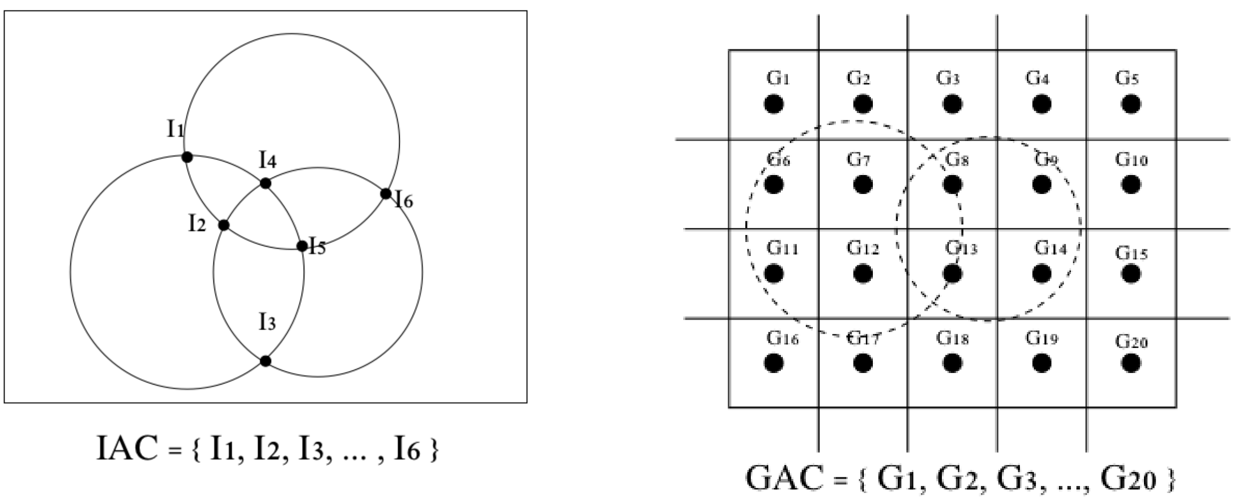
Unlike the *pure* coverage problems, LCRA problem needs to take SS's *SNR* requests into consideration, which makes the LCRA problem more complicated. For example, to solve pure coverage problems in [90], we could allocate circles' intersection points as candidate positions for coverage RSs, which are finite, to find best solutions. However, circles' intersection points might not guarantee feasible solutions for LCRA due to the *SNR* requirements. Though placing multiple RSs to cover multiple SSs can satisfy the distance requirements of SSs, it is likely that these RSs could interfere with each other, and result in unbearable *SNR* at some SSs.

To find a more appropriate solution for feasible coverage, we propose to use *small scale of grids spreading around entire field* as candidate RS locations. The benefit of using grids is that most of the field can be considered if we adjust the grid size small enough. However, smaller grid size will generate more candidate positions. Hence, the running time could be non-linearly increasing with smaller grid size. Therefore, *how to pick a right grid size to achieve the best tradeoff between solution quality and running time* is a critical issue. We propose two schemes to find best candidate positions:

1. *Intersections As Candidates* (IAC): *including all the intersection points between any two SS's feasible circles*, which is illustrated in Fig.2-2(a).

2. *Grids As Candidates* (GAC): *including all the center points of grids which divide the entire field*, which is shown in Fig.2-2(b).

It is easy to see that the number of candidate positions in GAC is highly related to the grid size. The smaller the grid size, the more accurate results we can obtain. Thus, we set the grid size as small as possible as long as optimizer software (e.g. Gurobi 5.0) can find results.



(a) IAC construction (b) GAC construction

Figure 2-2: Illustration of IAC and GAC

To solve LCRA, we first aim to find a minimum number of RS to cover SSs assuming RSs are using maximum transmit power (details in Section 2.1.4.1). Next step, in Section 2.1.4.2, we try to adjust transmit powers of RSs to further reduce the total power consumption without losing coverage. We outline the framework of our solution in the following. Details of each step are given in Section 2.1.4.1.

2.1.4.1 Coverage Under *SNR* Constraint

Given a relay network with a set of SS $S = \{s_1, s_2, s_3, ..., s_n\}$, corresponding feasible distance $D = \{d_1, d_2, d_3, ..., d_n\}$ and *SNR* threshold $\beta = \{\beta_1, \beta_2, \beta_3, ..., \beta_n\}$. We first formulate an

Algorithm 1: Framework of LCRA Solution

STEP_1: **Algorithm 2** - *Cover all the users by minimum number of RSs, assume each RS using its maximum power*
    STEP_1.1: **Algorithm 3** - *Split users into multiple zones, and then cover them zone by zone*
    STEP_1.2: **Algorithm 4** - *Try to identify users that are covered by only one RS*
    STEP_1.3: **Algorithm 5** - *Tweak RS positions to make them become feasible*
        STEP_1.3.1: **Algorithm 6** - *update coverage map once an RS's position is modified*
STEP_2: **Algorithm 7** - *Adjusting the power of each placed RS to further reduce energy consumption*

Integer Programming with quadratic constraints ILPQC to obtain optimal solutions. Let T_i and T_{ij} be the indicator variables in our ILPQC , where T_i indicates if candidate position i is chosen to place RS, and T_{ij} denotes if SS s_j has a feasible access link with RS at position i. The ILPQC is listed as below:

Objective

$$min \sum_{alli} T_i \tag{1.2}$$

Subject to :

$$T_i \leq \sum_{allj} T_{ij} \leq nT_i \qquad \forall i \tag{1.3}$$

$$\sum_{alli} T_{ij} = 1 \qquad \forall j \tag{1.4}$$

$$d_{ij}T_{ij} \leq d_j \qquad \forall i \forall j \tag{1.5}$$

$$\frac{d_{ij}^{-\alpha}}{\sum_{alli} d_{ij}^{-\alpha}T_i - d_{ij}^{-\alpha}T_i} \geq \beta_j T_{ij} \quad \forall j \tag{1.6}$$

where (1.2) is the objective to find the minimum number of RS positions. Linear constraint (1.3) states that each placed RS covers at least one SS. Linear constraint (1.4) states that each SS can access to only one RS. Linear constraint (1.5) states feasible distance requirement for each SS. Quadratic constraint (1.6) states that each SS should satisfy its *SNR* constraint. Both IAC and GAC are used to generate the set of candidate positions.

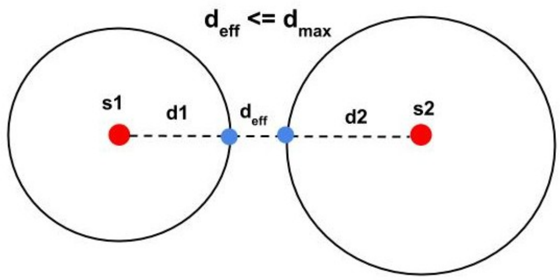
The formulation will provide the minimum number of RSs that can provide feasible coverage, and is used as the benchmark for performance evaluation in later sections. However, with the number of SSs increasing, the running time of the formulation with quadratic constraints increases exponentially. Therefore we propose a polynomial-time solution as a practical solution for large networks, which is listed in **Algorithm 2**.

Algorithm 2: *SNR* Aware Minimum Coverage (SAMC) (S,D,β)

Initialize set $L_{ss} = \{L_{ss}^1, L_{ss}^2, ..., L_{ss}^m\}$ which denotes SS groups to be returned from Zone Partition;
$L_{ss} \leftarrow$Zone Partition (S,D);
Initialize sets $L_{RS} = \{L_{RS}^1, L_{RS}^2, ..., L_{RS}^m\}$ which denotes each coverage RS group placed for each SS group;
for each SS group L_{ss}^i **do**
  $K_{mhs}^i = $ *Minimum Hitting Set* (L_{ss}^i, D_i);
  $G_i = $ Coverage Link Escape$(L_{ss}^i, D_i, K_{mhs}^i)$;
  $L_{RS}^i = $ Sliding Movement$(G_i, L_{ss}^i, D_i, \beta)$;
end for
for any L_{RS}^i in L_{RS} **do**
  if there exists a $L_{RS}^i = \emptyset$ **then**
    return *infeasible*;
  else
    $L_{RS} = L_{RS}^1 \bigcup L_{RS}^2 \bigcup ... \bigcup L_{RS}^m$;
    return L_{RS};
  end if
end for

The first step, Algorithm Zone partition, is to partition the field into several zones such that SSs and RSs in one zone will be distant from the stations in other zones. Thus, the interferences between *inter-zone* RS/SS pairs are small enough to be ignored. Details of Zone partition are presented in **Algorithm 3**.

Fig.2-3(a) and Fig.2-3(b) illustrate how the entire field can be divided into several independent zones. In Fig.2-3(a), the effective distance d_{eff} between s1 and s2 is calculated as $d_{eff} = dis(s1, s2) - d1 - d2$. If d_{eff} is less than or equal to d_{max}, which is the maximum distance between two SSs to ensure that one RS covering one SS may generate interferences to the other SS, we can add an edge from s1 to s2, which means that any RS placed to cover



(a) Effective distance

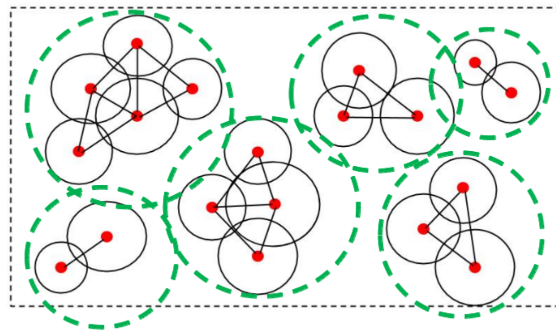

(b) Independent zones

Figure 2-3: Illustration of Zone partition

s1 may generate interferences to s2 (or vice versa). On the other hand, if d_{eff} is larger than d_{max}, s1 and s2 can be assigned to different independent zones. Any RS placed to cover s1 (or s2) will not interfere s2 (or s1). Using this scheme, we test each pair of SSs, and generate several independent zones as shown in Fig.2-3(b).

In *Line 5* of **Algorithm 2**, for SSs in each zone, we first find a set of RSs to cover all the SSs satisfying distance requirements by solving a *hitting set* problem. [62] proposes a $(1+\epsilon)$ approximation algorithm to solve minimum hitting set problem in geometry. Next, we aim to to satisfy the *SNR* requirements by adjusting RS positions. We notice that if one SS is covered by only one RS, named *one-on-one coverage*, then this RS could be moved closer to the covered SS (and hence further from other SSs). In this way, we can save power for the SS and RS over access links, and reduce the possibility of interfering other SSs. *Naturally, the more one-on-one coverage, the higher probability of satisfying SNR requirements for SSs.* To seek more one-on-one coverage, Coverage Link Escape (**Algorithm 4**) is used in *Line 6* of **Algorithm 2**.

After Coverage Link Escape, it is still possible that some RSs can only provide feasible distance coverage but not *SNR* for SSs. We call these place RS "*infeasible* RSs". To reduce *infeasible* RSs and improve the performance, in *Line 7* of **Algorithm 2**, we propose the Sliding Movement scheme, whose details are in **Algorithm 5**. For each infeasible RS location, which is on each covered SS's feasible circle, we try to "*slide*" the RS along the corresponding SS's feasible circle to try to find a feasible RS location. The question is *how to slide the infeasible candidates along SS's feasible circles*. The impact of sliding is complicated because it will not only *affect the signal power received by its covering SSs* but also *the noise received by other SSs*. One SS may receive higher *SNR* at the cost of other SSs suffering lower *SNR* as the result of a sliding operation. One method is to find infeasible coverage RSs which cannot satisfy *SNR* constraints. Then, based on the coverage topology, we try to slide the infeasible RSs along its covering feasible circles in order to clear *SNR* violations. If some *SNR* violation could not be cleared, we mark its covering RS as *un-slidable*. After sliding all infeasible RSs, we get a set of *slidable* RSs and their updated locations. Since updating slidable RSs can change the coverage topology, every SS' *SNR* constraint needs to be rechecked. To avoid of exponentially large number of updating of slidable RSs, we sort slideable RSs using the following criteria: $\frac{\textit{SNR gain for covered SS's after sliding}}{\textit{geneated noise to other SSs after sliding}}$, and slide all the violated RSs one by one in polynomial time. The details are in **Algorithm 6**. If all the SSs meet their *SNR* requests, we found a feasible solution for the SAMC problem. Otherwise, SAMC will return infeasible.

Algorithm 3: Zone Partition(S,D,N_{max})

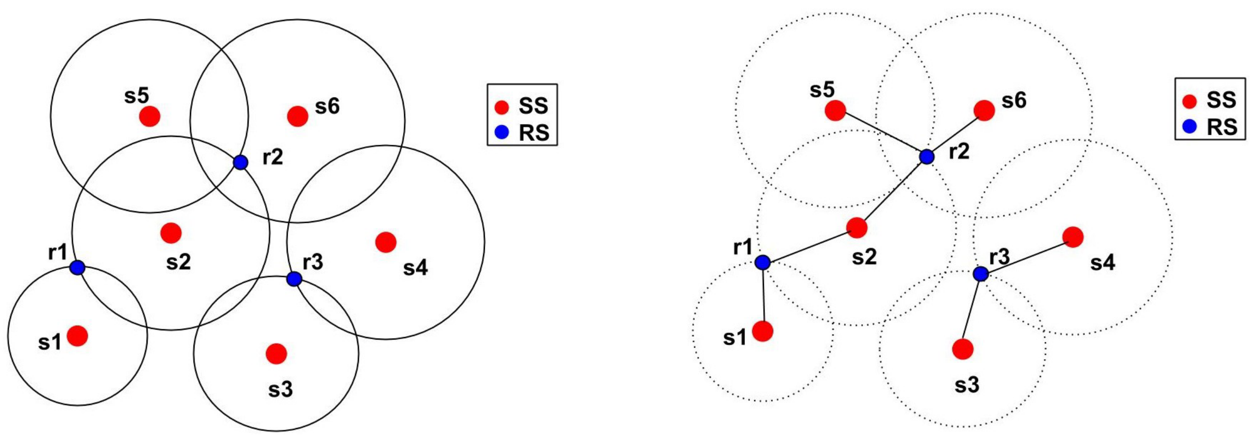
1: calculate d_{max} according to N_{max}, where $P_{max}Gd_{max}^{-\alpha} = N_{max}$, $G = G_tG_rh_t^2h_r^2$, N_{max} is the maximum noise which can be ignored;
2: create a new graph G involving all SSs in;
3: **for** any two SSs s_i, s_j in G **do**
4:     $d_{eff} \leftarrow \min\{\text{dist}(s_i,s_j)\text{-}d_i, \text{dist}(s_i,s_j)\text{-}d_j\}$;
5:     **if** $d_{eff} \le d_{max}$ **then**
6:         add edge $e(s_i, s_j)$ to G;
7:     **end if**
8: **end for**
9: find the *connected components* of G;
10: **return** SS groups of each connected component;

Let us use an example to illustrate how SAMC works. Because we divide the entire field into several independent zones, and all the operations in each independent zone are the same, we use one independent zone for our demonstration.

Fig.2-4(a) shows the results from minimum hitting set algorithm. There are 6 SSs in this independent zone. The best solution is placing 3 RSs to cover s1 to s6. In Fig.2-4(a), r1

Algorithm 4: Coverage Link Escape(S,D,K_{mhs})

construct a bipartite graph G between side A with all SSs, and side B including all the points in K_{mhs}, where K_{mhs} is the RS set returned by minimum hitting set algorithm;
for every SS s_i in side A **do**
  for every point p_i in K_{mhs} **do**
    if p_i is in or on c_i **then**
      add edge $e(s_i, p_i)$ to G;
    end if
  end for
end for
calculate $n_{max} \leftarrow$ the maximum number of edges including the same point in side B;
assume that all the edges in G and all the points in side B are not marked initially;
for n from n_{max} to 1 **do**
  for every unmarked point p in side B **do**
    if there are n edges containing p **then**
      mark these n edges;
      mark point p;
      for each recent marked edge $e(p, q)$ **do**
        delete all the unmarked edges containing point q;
      end for
    end if
  end for
end for
return bipartite graph G;



(a) Minimum hitting set

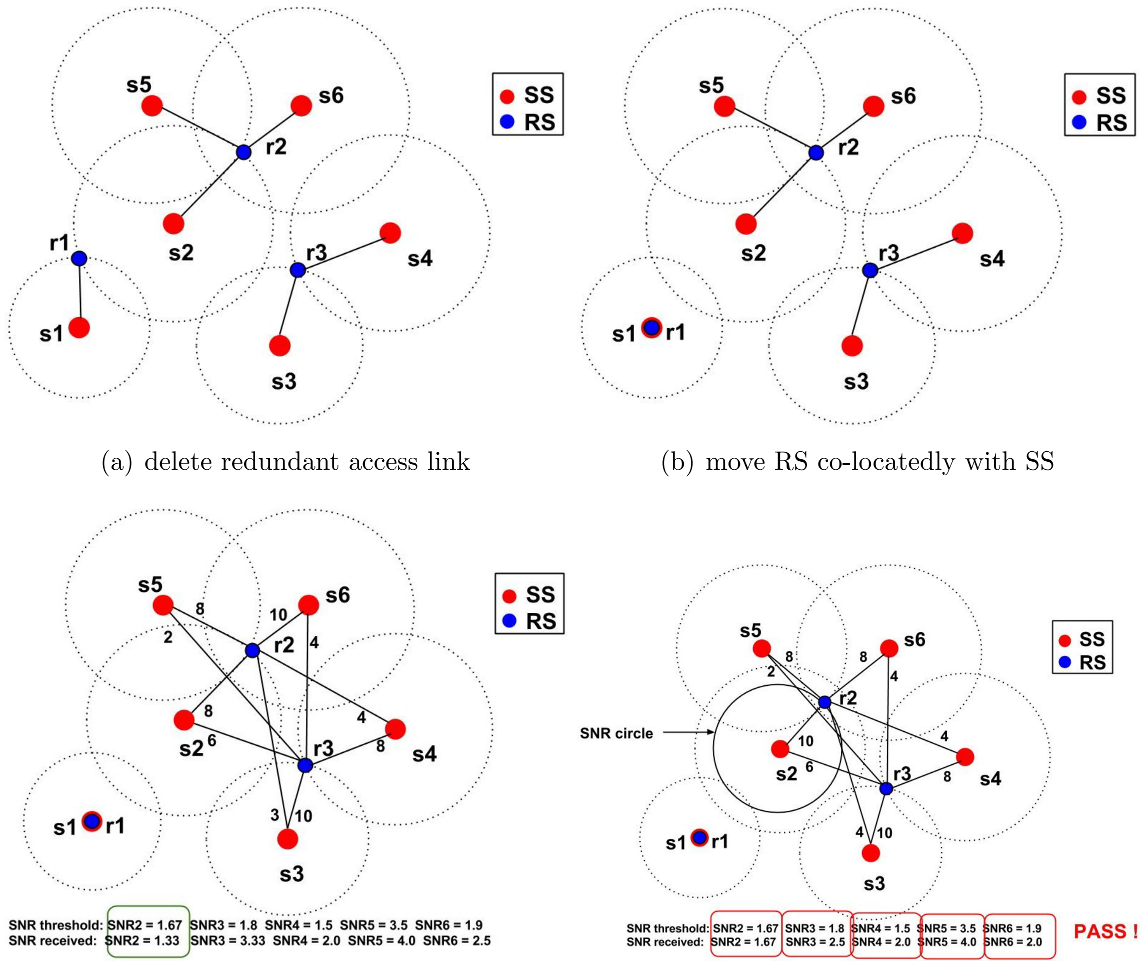
(b) Coverage link establishment

Figure 2-4: Minimum Hitting Set and Coverage Topology

covers s1 and s2, r2 covers s2, s5, s6, and r3 covers s3 and s4. Coverage links are established in Fig.2-4(b).

Fig.2-5(a) to Fig.2-5(d) show how Coverage Link Escape and RS Sliding Movement work, which are the core of SAMC. In Fig.2-5(a), as we know, one SS can only get access to one RS. In Fig.2-4(b) we see, s2 gets access to both r1 and r2. According to Coverage Link Escape scheme, the access link between s2 and r1 is a redundant link and it can be deleted since the degree of r1 is less than that of r2 ($2 < 3$). Now since r1 covers only one SS s1 after removing link (*s2, r1*), it can be moved to be co-located with s1 in order to avoid interfering with other SSs as Fig.2-5(b) shows. Since r1 co-locates with s1, it only needs a very low transmit power to maintain the coverage, and generate no interference to other SSs. Then we check each SS's *SNR* requirement in Fig.2-5(c), and find that s2's *SNR* requirement cannot be satisfied. In terms of RS Sliding Movement, r2 needs to be slided along s5's feasible circle in order to find a feasible location which is in the common area among s2's *SNR* circle, s5's feasible circle and s6's feasible circle. s2's *SNR* circle is shown in Fig.2-5(d). In Fig.2-5(d), after massaging r2's location, we recheck each user's *SNR* requirement. Eventually, SAMC finishes when all *SNR* requirements are satisfied.

In the beginning of SAMC algorithm, we invoke *minimum hitting set* algorithm to get the coverage RSs without considering *SNR* constraints. Then we are checking if each SS's *SNR* could be met using coverage RSs topology. If there exist some SSs whose *SNR* constraints are not satisfied, we need to slide coverage RS's point along its covering SSs' feasible circles in order to find a feasible solution. During the process of SAMC, *no coverage RSs are deleted or added in order to meet SSs' SNR constraints.* Consequently, *the result of SAMC has the same number of coverage RSs as the number returned from the minimum hitting set solution.* Therefore, SAMC's performance is highly related to minimum hitting set algorithm, following the same scheme used in [90]. [62] gives an $(1 + \epsilon)$-approximation PTAS for the minimum hitting set problem. We adopt the PTAS, and claim that if SAMC returns a feasible solution, it is also an $(1+\epsilon)$-approximation solution. In other words, if SAMC returns a feasible solution K, the number of RS provided by K will be no more than $(1 + \epsilon) * |OPT_C|$, where OPT_C is an optimal solution with the minimum number of RSs that can provide feasible coverage.



(a) delete redundant access link

(b) move RS co-locatedly with SS

(c) test *SNR* requirements

(d) slide violated RS

Figure 2-5: Illustration of Coverage Link Escape and RS Sliding Movement

2.1.4.2 Power Reduction Optimization

In the previous section, we find feasible coverage RSs assuming that RSs are transmitting at their *maximum* powers in SAMC. In this section, we aim to *adjust transmit powers* of the placed RSs so that we can *further reduce the energy consumption while maintain coverage and SNR constraints.*

Given a fixed network topology consisting of SSs and *coverage* RSs found by SAMC, we first present another Linear Programming with Quadratic Constraints (LPQC) to get an optimal RS transmit power allocation so that the total transmit power is minimized. Let P_i denote the transmit power of i^{th} RS in coverage RSs, which is in the range of $[0, P_{max}]$, and

Algorithm 5: RS Sliding Movement(G, S, D, β)

$H \leftarrow \emptyset, B \leftarrow \emptyset$;
for every point p in side B of G **do**
  if there is only one edge $e(p, q)$ containing p **then**
    if p and q are not at the same location **then**
      move p to the same location as q;
    end if
    $H \leftarrow H \bigcup \{p\}$;
    delete point p and corresponding SS in G;
  end if
end for
for every SS s_i in side A **do**
  check if SNR constraint β of s_i can be satisfied
  if not, mark s_i;
end for
$B = B \bigcup \{\text{all marked } s_i\}$;
if B is empty **then**
  $H = H \bigcup \{\text{all RSs in side B}\}$;
  return H;
else
  $R_u^s \leftarrow$ all the RSs in side B covering the SSs in B;
  $R_r^s \leftarrow$ all the rest RSs in side B;
  $H' \leftarrow$ update RS topology $(R_u^s, R_r^s, G, S, D, H, B)$;
  if $H == H'$ **then**
    $H' \leftarrow \emptyset$;
  end if
  return H'
end if

T_{ij} be the indicator of whether SS s_j communicates with RS_i. The goal is to $min \sum P_i$.

Objective :

$$min \sum_{alli} P_i \tag{1.7}$$

Subject to :

$$\sum_{alli} T_{ij} = 1 \qquad \forall j \tag{1.8}$$

$$P_i G d_{ij}^{-\alpha} \geq P_{ss}^j T_{ij} \qquad \forall j \tag{1.9}$$

$$\frac{P_i G d_{ij}^{-\alpha}}{\sum_{alli} P_i G d_{ij}^{-\alpha} - P_i G d_{ij}^{-\alpha}} \geq \beta_j T_{ij} \quad \forall j \tag{1.10}$$

According to the two-ray model, $P_r = P_t G_t G_r h_t^2 h_r^2 d^{-\alpha}$, and $G = G_t G_r h_t^2 h_r^2$ are all constants. Constraint (1.8) means that any SS must communicate with one and only one RS. Constraint (1.9) indicates that coverage RS r_i must provide enough transmit power to ensure the data

Algorithm 6: Update RS Topology $(R_u^s, R_r^s, G, S, D, H, B)$

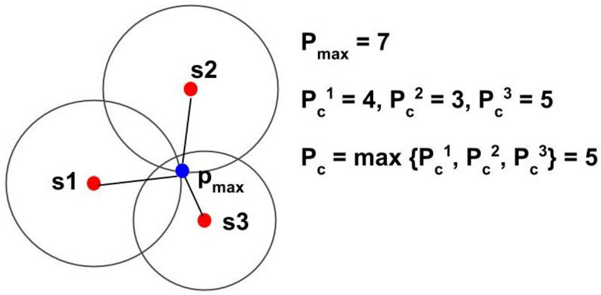
for each RS r_i in R_u^s **do**
  $W \leftarrow \emptyset$;
  let s_k and s_j denote SSs whose SNR can and cannot be met under coverage of r_i, respectively;
  construct a virtual circle c_j' for each s_j to ensure that $s_j's$ SNR can be met only if r_i moves into c_j';
  W=$W \bigcup$ {all virtual circles c_j'} $\bigcup$ {all feasible circles c_k of s_k};
  if all the circles in W have common area **then**
    mark r_i as slidable to r_i' in R_u^s, where r_i' is the centre of the common area;
  else
    mark r_i as un-slidable in R_u^s;
  end if
end for
for each slidable r_i in R_u^s **do**
  let s_j be such that s_j's SNR cannot be satisfied under the coverage of r_i;
  $\triangle_{snr}^i \leftarrow SNR_{r_i'}^{s_j} - SNR_{r_i}^{s_j}$;
  $S_i \leftarrow S/s_j$; $I_i \leftarrow 0$;
  for each SS s_k in S_i **do**
    $I_i^k \leftarrow I_{r_i'}^{s_k} - I_{r_i}^{s_k}$;
    $I_i \leftarrow I_i + I_i^k$;
  end for
  $\triangle_i \leftarrow \frac{\triangle_{snr}^i}{I_i}$;
end for
construct the pairs $(\triangle_i, r_i, r_i')$;
$i_{max} \leftarrow argmax_{i \in |R_u^s|} \triangle_i$;
update $r_{i_{max}}$ to $r_{i_{max}}'$ and obtain an updated R_u^s;
if all SNRs satisfied **then**
  $H \leftarrow H \bigcup R_r^s \bigcup R_u^s$;
else
  record the unsatisfied SSs into a new set B';
  if size(B') < size(B) **then**
    $R_u^{s'} \leftarrow$
    all the RSs in side B covering the SSs in B';
    $R_r^{s'} \leftarrow$ all the rest RSs in side B;
    $H \leftarrow$
    Update RS Topology $(R_u^{s'}, R_r^{s'}, G, S, D, H, B')$;
  end if
end if
return H;

rate request from its covering SS s_j, where P_{ss}^j denotes minimum received power requested by SS s_j. Quadratic constraint (1.10) represents the SNR constraint for every SS. In numerical results, we will use the LPQC as the optimal solution for power cost reduction and the benchmark for performance comparison. Similarly, LPQC takes exponentially increased running time as the number of RSs or SSs increasing, it is not efficient or practically usable for large networks. Therefore, we present another efficient heuristic based on the following observation.

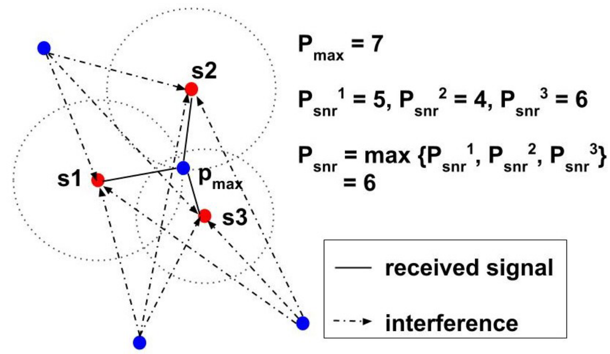
We observe that the reduction of transmit power of an RS will reduce the noise to SSs covered by other RSs so that these SSs could have higher probability to fulfil their *SNR* constraints. We call *the minimum transmit power of an RS under its coverage constraints* P_c. Besides coverage constraints, RSs need to meet each SS's *SNR* constraint. Similarly, we call *the minimum transmit power of RS under its SNR constraints* P_{snr}. *As long as the transmit power of one RS is no less than* P_c *and* P_{snr} *for its covered SSs, its transmit power can be reduced.* Let L_{low} be the coverage RS set as a result of SAMC, β be the *SNR* threshold, P_{max} be the maximum transmit power of RS, and P_{ss} be the set of minimum received power each SS needs to ensure its data rate. Moreover, let P^i_{min} denote the coverage power P_c for RS r_i and P^i_{SNR} denote the *SNR* power P_{snr} for RS r_i, respectively. It is straightforward to calculate coverage power P_c and *SNR* power P_{snr} for each RS r_i. If all the RSs can reduce power to their own coverage power P_c while meeting *SNR* constraints, the power saving approach is optimal.

The details of the power saving algorithms are listed in **Algorithm 7**. And let us use an example to illustrate how to calculate the coverage power P_c and the *SNR* power P_{snr} for each RS.

In Fig.2-6(a), one RS covers s1, s2 and s3. Without taking *SNR* into consideration, we can figure out the minimum power to cover s1, s2 and s3, respectively. This example shows the maximum transmit power of RS is 7, the minimum power to cover s1, s2 and s3 are $4, 3, 5$, respectively. Then the coverage power p_c for this RS is the maximum value among p^1_c, p^2_c and p^3_c, which is 5. In order to calculate *SNR* power for one RS, we need to take received interference to each SS under the coverage of the RS into consideration. In Fig.2-6(b), the received interference by s1, s2 and s3 can be easily determined since we assume that three



(a) Coverage power



(b) *SNR* power

Figure 2-6: Coverage Power and *SNR* power calculation

surrounding RSs contributing to interference on s1, s2, s3 are transmitting in their maximum transmit power. Then we can reduce the transmit power of center RS from its maximum value to satisfy s1's, s2's and s3's *SNR* requirements, respectively. We obtain $p^1_{snr} = 3$, $p^2_{snr} = 2$ and $p^3_{snr} = 6$. The *SNR* power for center RS can be easily calculated by taking the maximum value among $p^1_{snr}, p^2_{snr}, p^3_{snr}$ since all three SSs under coverage must satisfy their *SNR* requirements simultaneously.

Theorem 1. Algorithm 7 is a $(1+\phi)$-approximation for the Power Reduction Optimization (PRO) problem. More specifically, if the power cost of all the RSs returned by Algorithm 7 is denoted by $|P|$, we have $|P| \leq (1+\phi) \cdot |OPT_P|$, where $|OPT_P|$ is an optimal solution for PRO, and $\phi = \frac{\sum_{i \in \mathcal{C}}(P^i_{snr} - P^i_c)}{OPT_P}$. □

Proof: If all $P_{snr} \leq P_c$, then $|P| = |OPT_P|$. Otherwise, let P^i_c denote the coverage power for RS r_i, and P^i_{snr} denote the *SNR* power for RS r_i. Thus in whatever OPT_P or P, it is composed of P^i_c or P^i_{snr} for each RS r_i. For instance,

$$P = \{P^1_c, P^2_{snr}, P^3_{snr}, P^4_{snr}, P^5_c\}$$

$$OPT_P = \{P^1_c, P^2_c, P^3_{snr}, P^4_{snr}, P^5_c\}$$

Also, we let $I = \max_i\{\text{all } P^i_{snr} \text{ occur in } OPT_P\}$ and C be the set of r_i for all $i \in [1, I]$ in OPT_P which does not operate in P^i_{snr}.

Algorithm 7: *Power Reduction Optimization (PRO)* $(L_{low}, S, P_{ss}, \beta, P_{max})$

$K \leftarrow \emptyset; P_1 \leftarrow \emptyset; P_2 \leftarrow \emptyset; P_3 \leftarrow \emptyset; P_{tmp} \leftarrow \emptyset;$
Initialize P_1, P_2, P_3, P_{tmp}
for each item i in L_{low} **do**
    $P_1^i = P_{max}$; $P_3^i = P_{max}$; compute P_{min}^i;
    $P_2^i = P_{min}^i$; $P_{tmp}^i = P_{max}$;
end for
put each RS point of L_{low} into K;
while K is not empty **do**
    for each item i in P_1 **do**
        if $P_1^i == P_3^i$ **then**
            $P_1^i = P_2^i$;
            check if P_1^i can meet SNR constraints for SS covered by RS_i;
            if yes **then**
                remove RS point i from K; $P_{tmp}^i = P_1^i$;
            end if
            $P_1^i = P_{max}$;
        end if
    end for
    clear P_1; $P_1 \leftarrow P_{tmp}$;
    if length of K is not changed **then**
        for each item i in P_1 **do**
            if $P_1^i == P_3^i$ **then**
                compute P_{SNR}^i;
            end if
        end for
        Find index i for minimum $\Delta P_i = P_{SNR}^i - P_{min}^i$;
        $P_1^i = P_{SNR}^i$; $P_{tmp}^i = P_{SNR}^i$;
        remove RS point i from K;
    end if
end while
return $\sum_{alli} P_1^i$;

Therefore, the worse case for P is,

$$P = OPT_P + \sum_{i \in C} (P_{snr}^i - P_c^i)$$

The approximation ratio in worse case is

$$\frac{P}{OPT_P} = \frac{OPT_P + \sum_{i \in C}(P_{snr}^i - P_c^i)}{OPT_P}$$

$$= 1 + \frac{\sum_{i \in C}(P_{snr}^i - P_c^i)}{OPT_P}$$

Since $\phi = \frac{\sum_{i \in C}(P_{snr}^i - P_c^i)}{OPT_P}$, we have $|P| \leq (1 + \phi) \cdot |OPT_P|$. ■

2.1.5 Approximation Solutions for UCRA

After covering all the SSs with sufficient SNR, we need to relay the traffic from the covering RSs to the BSs. In [90], the authors studied a similar MUST problem, which is estimated to be *NP*-hard. MUST assumes only one BS and RSs always operate with maximum power. Therefore, MUST can be regarded as a special case of UCRA. To solve UCRA, the first challenge is how to decide the feasible distance of each RS, which is affected by the SSs or RSs being covered. In order to guarantee the data rate of each SS, for each RS r_i, the link capacity between r_i and its parent node cannot be lower than the one between r_i and its any child. Therefore, we define *feasible distance of an connectivity RS r_i, connecting r_i and its parent station (an RS or a BS), should equals to the minimum feasible distance of all its children*, which is shown in Fig. 2-7.

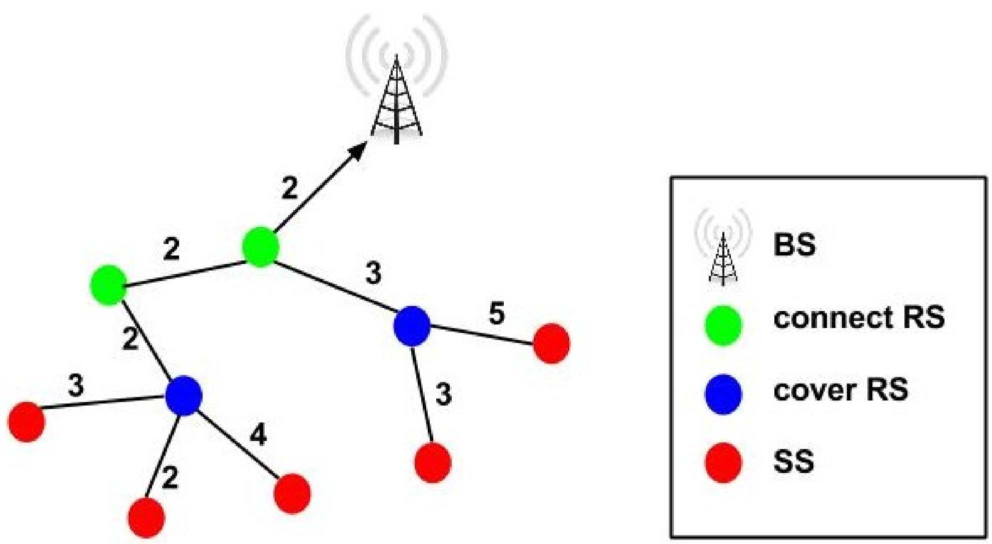


Figure 2-7: Feasible distances of connectivity RSs

With the assumption of connectivity RSs operating with P_{max}, we propose our solution in **Algorithm 8**.

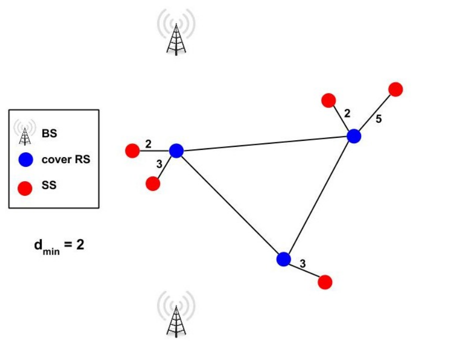
Let us use an example to demonstrate Algorithm MBMC, which is shown in Fig. 2-8(a) to Fig. 2-8(e). In this example, there are 3 BSs, 5 SSs and 3 coverage RSs deployed in this field. Firstly, we calculate the overall minimum feasible distance, which is 2 in Fig. 2-8(a).

Algorithm 8: Multiple BS Minimum Connectivity (MBMC)(R_c, S, D, B)

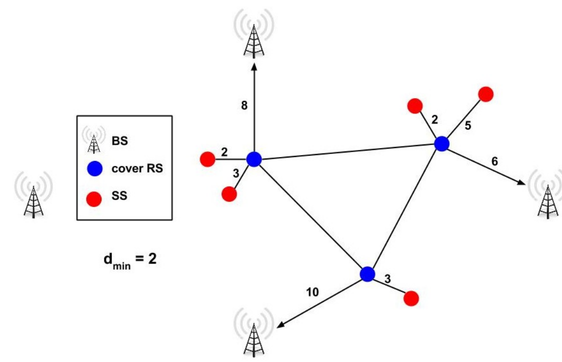
construct a complete graph $G = (R_c, E)$, where R_c denotes coverage RS set;
$d_{min} = \min_{i \in S} d_i$;
for each node r_i in G **do**
  create a new set K_i;
  for each BS b_j in B **do**
    calculate distance (r_i, b_j) and store it into K_i;
  end for
  find $\min K_i$ and add the corresponding BS node b into G;
  add edge $e(r_i, b)$ into G;
end for
for each edge $e(x_i, x_j)$ in G **do**
  assign weight $w_1(x_i, x_j) = \lceil \frac{\|e(x_i,x_j)\|}{d_min} \rceil - 1$ on the edge;
end for
Find a minimum spanning tree $\tau_m st$ of G with BS as the root;
for each RS r_i **do**
  Calculate its feasible distance d_r^i;
end for
for each RS r_i and its parent r_i^p on $\tau_m st$ **do**
  $w_2(r_i^p, r_i) = \lceil \frac{\|e(r_i^p, r_i)\|}{d_r^i} \rceil - 1$;
  Place $w_2(r_i^p, r_i)$ RSs on $e(r_i^p, r_i)$ separating the edge into $\lceil \frac{\|e(r_i^p, r_i)\|}{d_r^i} \rceil$ sections with each one with feasible distance;
end for

In Fig.2-8(b), each coverage RS builds an edge to its nearest BS. The numbers on the edges are the distances between coverage RSs to their nearest BS. Then we add additional edges to build a complete graph among all the coverage RSs. The weights on all the edges are calculated according to the scheme presented in MBMC. All the edge weights are shown in Fig.2-8(c). Since all the BSs are backhauled to a central location, we assume that there is an edge with weight of 0 between any two BSs. Then we find the minimum spanning tree of this graph and obtain the results shown in Fig.2-8(d). Finally, connectivity RSs are placed on each edge between coverage RS and its connecting BS by equally separating the edge. Fig.2-8(e) shows the results of MBMC.

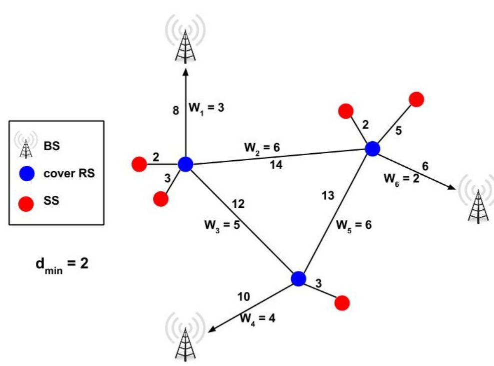
Since both MBMC and MUST proposed in [90] are minimum spanning tree based algorithms, MBMC has the same $\frac{8d_{max}}{d_{min}}$-approximation ratio as MUST proved, where d_{min} and d_{max} denote the minimum and maximum feasible distances of SSs, respectively. Having locations



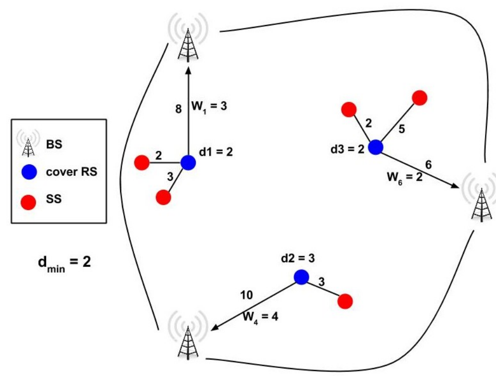
(a) calculate overall d_{min}



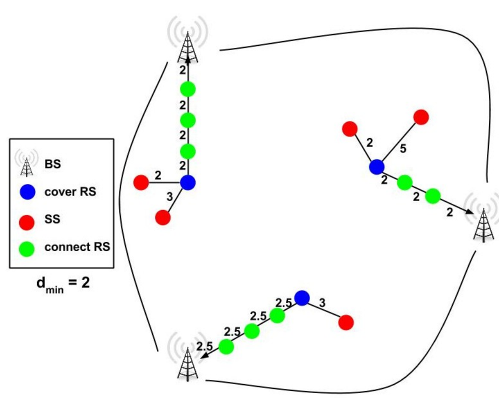
(b) select nearest BS



(c) calculate link weights



(d) minimum spanning tree



(e) deploy connectivity RSs

Figure 2-8: Illustration of MBMC

of connectivity RSs returned by MBMC, we then try to optimize power cost of each RS. Our solution is listed in **Algorithm 9**. Let L_{low} denote the set of *coverage* RSs, L_{high} denote the set of *connectivity* RSs, P_{ss}^{i} denote the set of received power requirements of SSs covered by RS r_i, P_{rs}^{i} denote received power requirement of RS r_i, N_i denote the number of RSs placed on the path from RS r_i to its parent, p_{ij} denote the transmit power of j^{th} RS on the path from RS r_i to its parent, and $G = G_t G_r h_t^2 h_r^2$.

Algorithm 9: Upper-tier Connectivity Power Optimization (UCPO)$(L_{low}, L_{high}, P_{ss})$

```
for each RS r_i in L_low do
    put each P^i_ss into new set K_i;
    P^i_rs = max(K_i);
    D_i = distance(i,parent(i)) / N_i;
    P_i = P^i_rs / (G D_i^{-α});
    for each RS r_j on path (i, parent(i)) do
        p_ij = P_i;
    end for
end for
return Σ_alli Σ_allj p_ij;
```

2.1.6 Approximation Algorithm for SAG problem

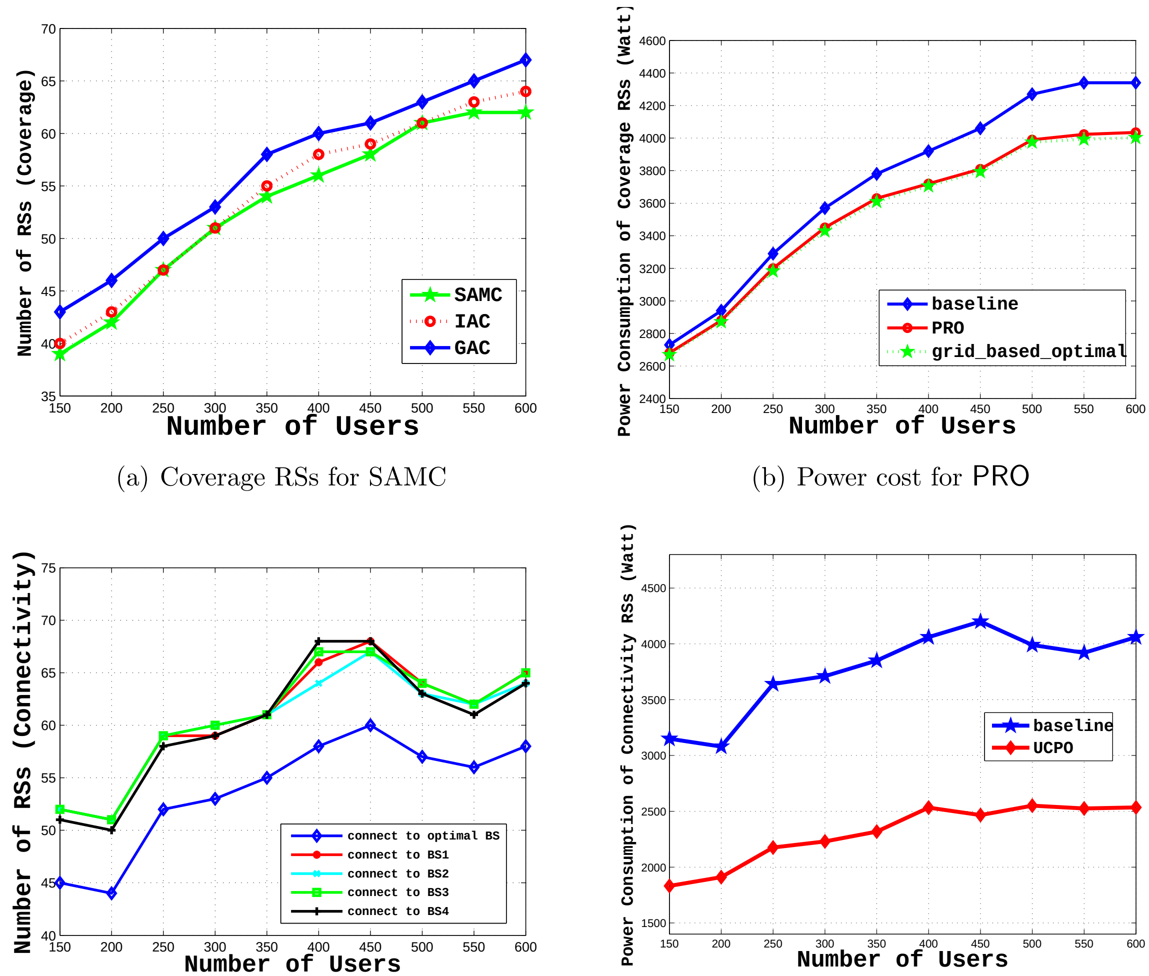
With the approximation solutions (in terms of number of RSs placed) for both lower tier and upper tier, we present an approximation algorithm for the SAG problem in **Algorithm 10**.

Algorithm 10: *SNR*-aware Green (SAG) Relay$(S, D, B, \beta, P_{ss}, P_{max})$

```
L_low ← ∅; L_high ← ∅;
L_low ← SAMC(S, D, β);
P_L ← PRO(L_low, S, P_ss, β, P_max);
L_high ← MBMC(L_low, S, D, B);
P_H ← UCPO(L_low, L_high, P_ss);
P_total = P_L + P_H;
return P_total;
```

2.1.7 Numerical Results

In this section, numerical results are presented to show the effectiveness of our schemes, including SAMC, PRO, MBMC, UCPO and SAG algorithms. All the simulations are run on a Intel Core(TM) i5 CPU of 2.7GHz with 8GB memory. All the SSs and BSs are uniformly distributed in a square testing field. All the figures illustrate the average of 10 test runs for various scenarios.



(a) Coverage RSs for SAMC

(b) Power cost for PRO

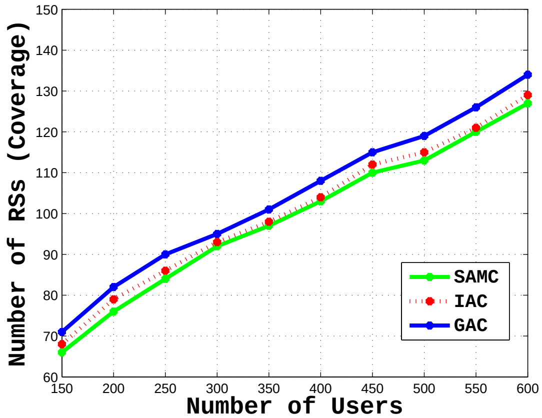
(c) Connectivity RSs for MBMC

(d) Power cost for UCPO

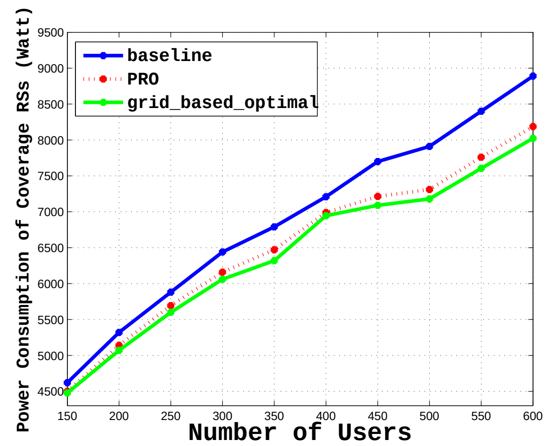
Figure 2-9: Performance in $3km \times 3km$ playing field

2.1.7.1 Simulation Environment Settings

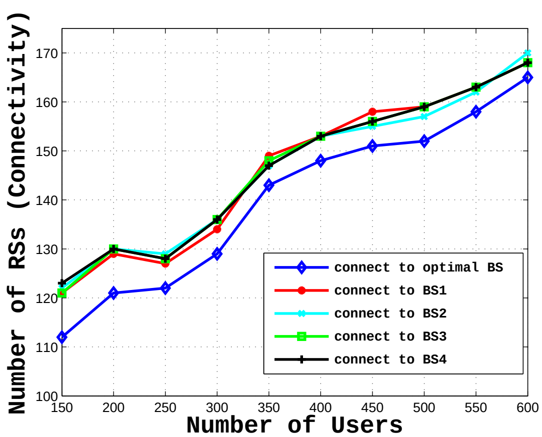
Since solving the ILP with quadratic constraints in Gurobi 5.0 [36] takes exponentially increasing running time and memory as growing the number of SSs or lessening the grid size, very large scale of testing field and huge amount of SSs are not considered in our simulations. We consider the large scale of playing field is composed of a couple of small fields and the operations in each sub-field are independent to others. More specially, the entire testing field can be divided into several sub-zones depending on the distributions of SSs in Zone Partition Algorithm. We select two scales of testing field for our numerical evaluations: $3km \times 3km$ field and $5km \times 5km$ field. And we set the grid size as small as possible as long as we can



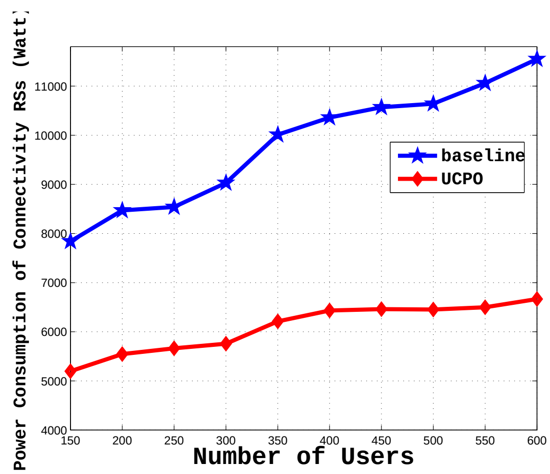
(a) Coverage RSs for SAMC



(b) Power cost for PRO



(c) Connectivity RSs for MBMC



(d) Power cost for UCPO

Figure 2-10: Performance in $5km \times 5km$ playing field

avoid out-of-memory issue from solving our ILPs. Signal-to-Noise Ratio (SNR) threshold for each SS is set according to typical 802.16 standard document. In 802.16 standard, each user with a certain data rate request needs to satisfy a minimum (SNR) threshold requirement. Data rate request for each user is randomly distributed between $10Mb/s$ and $45Mb/s$. The number of SSs in playing fields varies from 150 to 600, which are uniformly distributed as well. We place at most 4 BSs in the testing field in order to show the performance of MBMC comparing to MUST in previous literature. Now, we have five metrics to be compared among various scenarios such as *the number of coverage RSs, power consumption of coverage RSs, the number of connectivity RSs, power consumption of connectivity RSs* and *the entire power* *consumption of all relay nodes*. First, we present the numerical results on both lower tier and upper tier, separately. Then we show the performance of our SAG scheme comparing with some other existing schemes. The results collected from lower tier and upper tier are shown in 2.1.7.2 and 2.1.7.3, respectively.

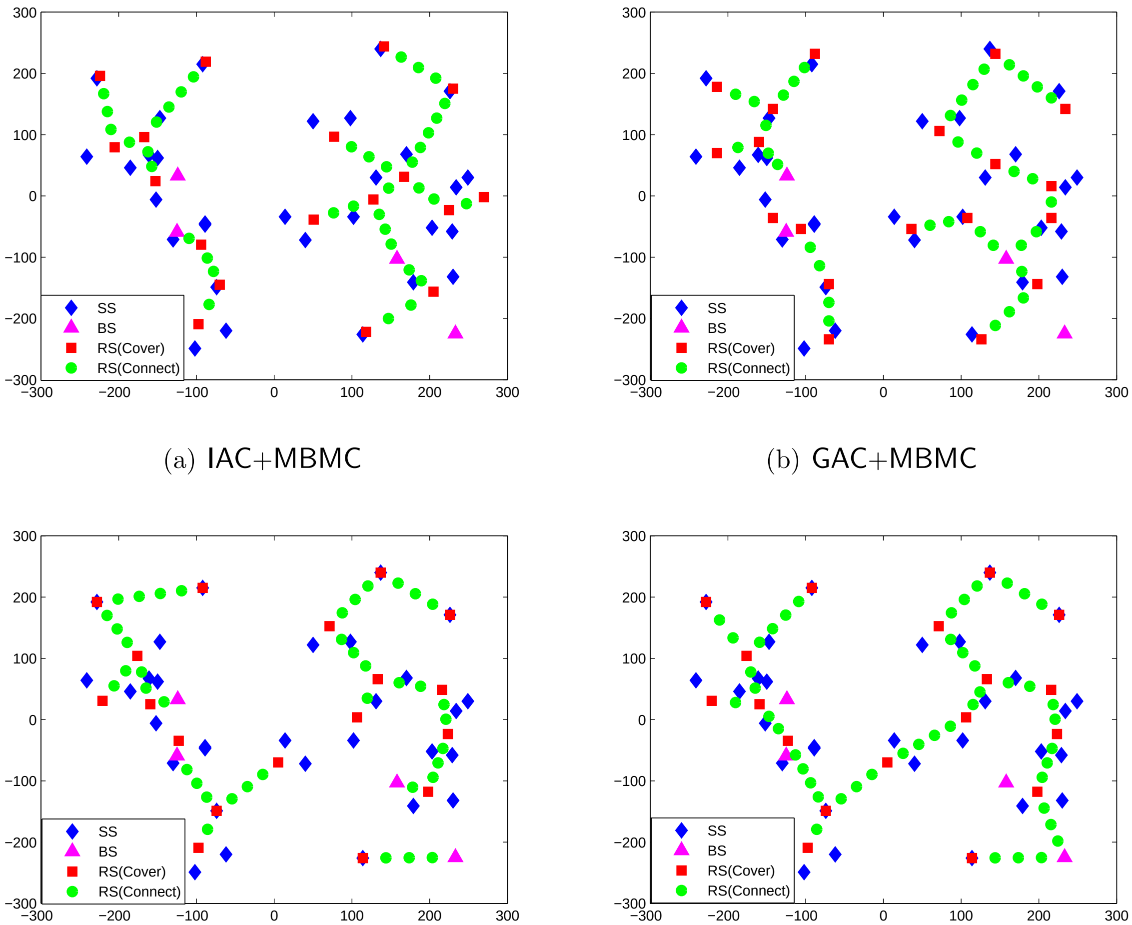
Here is a list of all constant parameters we used in our simulation.

Parameters	**Values**
Max. Transmit Power of RS	70 Watt (48.45 dBm)
Channel Bandwidth	10 MHz
Height of User Client	1.5 m
Height of RS	10 m
Transmitter Antenna Gain	2 dBi
Receiver Antenna Gain	2 dBi
Attenuation Factor	2
Thermal Noise	-85 dBm
Grid Size	100 m$\times$ 100 m

Table 2.2: Constant Parameters

2.1.7.2 Evaluation of Heuristics on Lower Tier

On the lower tier, we test the performance of IAC, GAC and SAMC on two playing fields of $3km \times 3km$ and $5km \times 5km$, respectively. The results are shown in Fig. 2-9(a) and Fig. 2-10(a). We can easily see that the number of coverage RSs coming from SAMC is lower than both GAC and IAC in whichever scenario. GAC has the most number of coverage RSs, which is caused by the selected size of candidate grid. The less size of candidate grid, the more accurate the results it would provide. Due to limited amount of memory in our simulation computer, we are not able to set small enough grid size in order to get the near-optimal solution so that the results from GAC are not as good as the results from IAC. Our proposed SAMC is starting from the results of minimum hitting set based on IAC. If the selected locations of coverage RSs from minimum hitting set can not satisfy all SSs' SNR threshold requirements, SAMC tries to slide violated RS location along the feasible circle of SS in order to find a location that can satisfy previous violated SNR threshold requirement.



(a) IAC+MBMC (b) GAC+MBMC

(c) SAMC+MBMC (d) SAMC+MUST

Figure 2-11: Illustration of tree topologies for various schemes

But IAC based ILPQC will not perform these following improvements. IAC based ILPQC just drops this RS location and then searches one or more RSs to replace the RS. Since the algorithm we select for solving minimum hitting set problem is a near-optimal solution, it is probable that IAC based ILPQC would find more than one RSs to replace one violated RS while ensuring the SSs under the coverage of the violated RS can still be covered. If there are many violated RSs which are selected from minimum hitting set, it is likely that IAC based ILPQC would return more number of coverage RSs than SAMC returns. Such possibility is verified by our results in Fig. 2-9(a) and Fig. 2-10(a). When all the users' SNR requirements are satisfied based on the results from minimum hitting set, IAC based ILPQC will probably return the same amount of coverage RSs as SAMC does, which can also be seen in both Fig. 2-9(a) and Fig. 2-10(a). From above observations, we can see that SAMC outperforms both IAC and GAC in terms of not only the number of coverage RSs but also the running times.

Fig.2-9(b) and Fig.2-10(b) show that PRO performs near to optimal and does save much power from the *baseline* model, in which all the RSs operate in *maximum power* when more users are involved. Moreover, PRO can save more power comparing with the baseline especially in larger scale of testing field with same set of users uniformly distributed. Therefore, this result confirms our theoretical analysis of PRO performance.

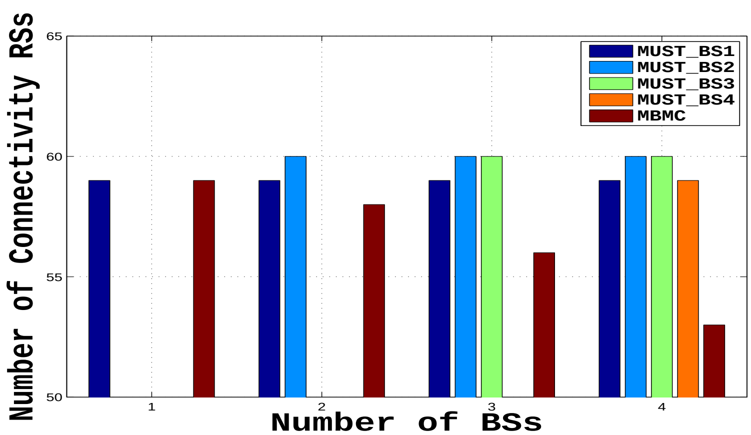


Figure 2-12: Compare the performance between MBMC and MUST with various number of BSs in $3km \times 3km$ field($N_{SS} = 300$)

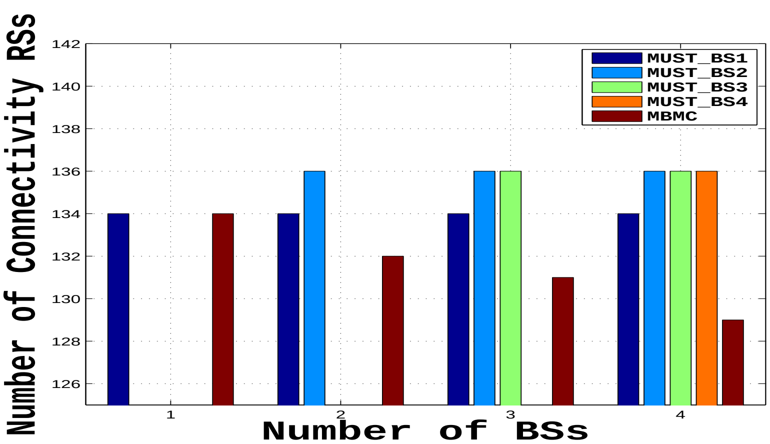


Figure 2-13: Compare the performance between MBMC and MUST with various number of BSs in $5km \times 5km$ field($N_{SS} = 300$)

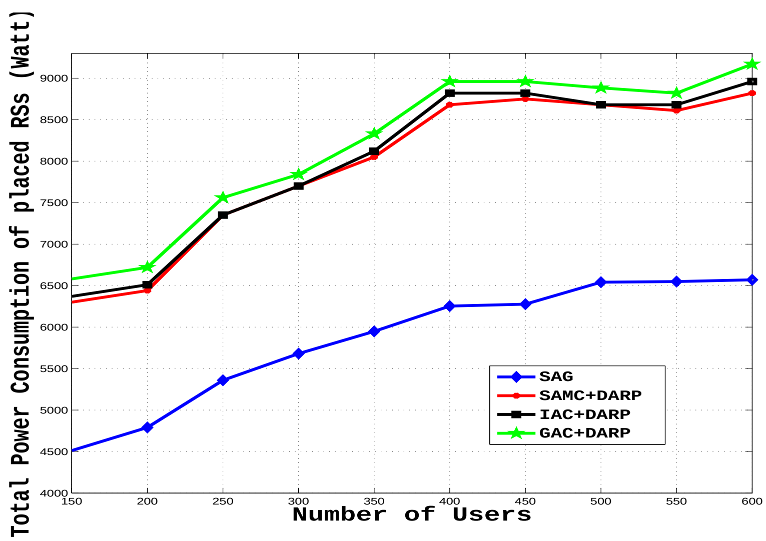


Figure 2-14: SAG performance in $3km \times 3km$ field

2.1.7.3 Evaluation of Heuristics on Upper Tier

On the upper tier, we are concentrating on showing how MBMC works and why it outperforms MUST proposed in [90]. As we discussed in previous section, MUST can only be applied to *one BS scenario* but MBMC extents MUST scheme and works well in *multiple BSs* environment, which is the more practical deployment. Thus we claim that MBMC is more practical than MUST. Assume that 4 BSs are deployed in the testing field. We run MUST for four times, for each of which we let MUST connect to one of the four BSs, respectively. Fig. 2-11(d) illustrates the case in which all SSs only connect to the corner BS (MUST algorithm) and Fig. 2-11(c) illustrates the case in which all SSs connect to their nearest BS (MBMC algorithm). We can compare the data collected in Fig. 2-9(c) and Fig. 2-10(c) between MBMC and MUST. Apparently, MBMC outperforms MUST from each of the scenarios adapting MUST. Also, we test the scenarios of various number of BSs from 1 to 4 on both $3km \times 3km$ and $5km \times 5km$ playing fields with the number of users set to 300. From Fig. 2-12 and Fig. 2-13, we can easily find that the number of connectivity RSs from MBMC is less than or equal to that from MUST in all testing scenarios. If there is only one BS deployed in the field, MBMC and MUST return the same result. However, when the number of BSs is

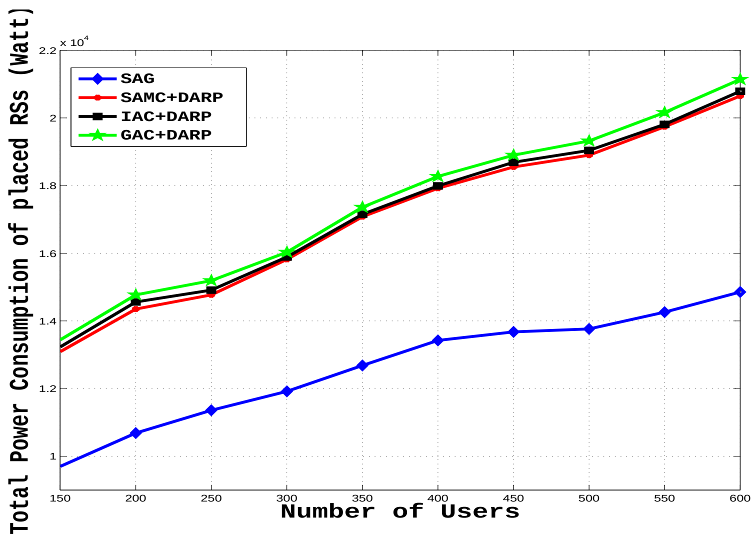
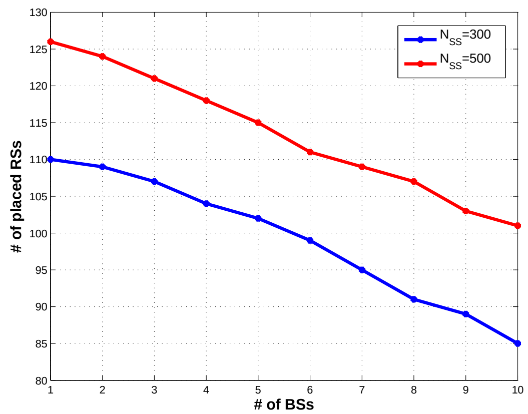


Figure 2-15: SAG performance in $5km \times 5km$ field

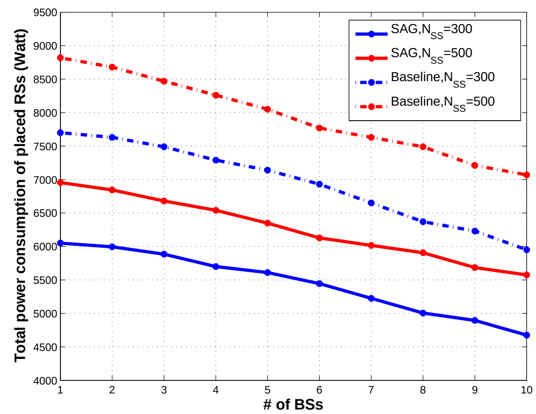
increasing from 2 to 4, we can see from both Fig. 2-12 and Fig. 2-13 that the number of connectivity RSs returned from MBMC is decreasing because there are more BSs to connect to and each coverage RS will choose its nearest BS to connect to according to MBMC, which leads to less number of connectivity RSs to place. From the above observation, we can say that MBMC outperforms MUST in more practical environment. Based on the connectivity topology returned by MBMC, it is probably not necessary for each connectivity RS to transmit in its maximum transmit power to maintain the connection. We reduce the transmit power of each connectivity RS according to UCPO scheme and then find that large amount of power consumption can be saved comparing with the baseline in which all connectivity RSs are transmitting in maximum power. The performance of optimal UCPO can be confirmed in Fig. 2-9(d) and Fig. 2-10(d).

2.1.7.4 Evaluation of Heuristics for SAG

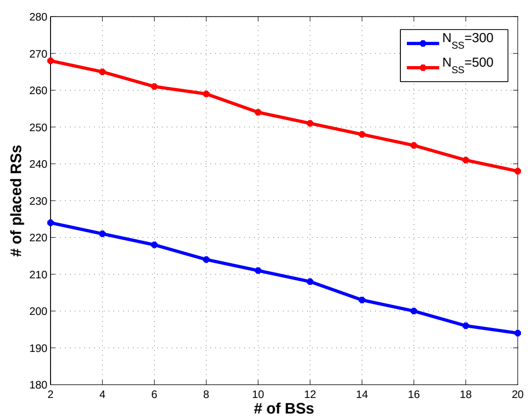
Our SAG scheme combines the solutions for both lower tier and upper tier. Fig. 2-11(a), Fig. 2-11(b) and Fig. 2-11(c) illustrate the tree topologies coming from IAC plus MBMC, GAC plus MBMC and SAMC plus MBMC, respectively. At last, we compare the performance among SAG, SAMC+DARP, IAC+DARP and GAC+DARP, where DARP represents the deployment



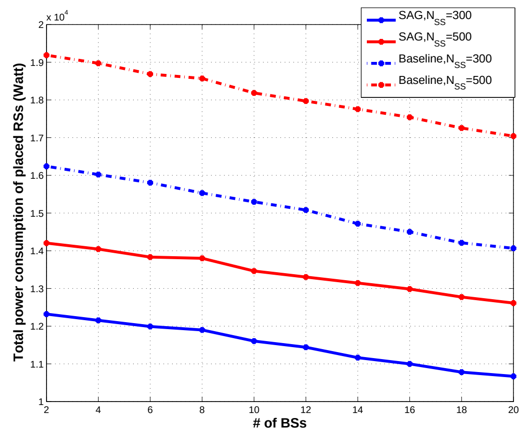
(a) $3km \times 3km$ field



(b) $3km \times 3km$ field



(c) $5km \times 5km$ field



(d) $5km \times 5km$ field

Figure 2-16: Impact of the number of BSs

approaches proposed in [90] excluding their lower tier coverage approaches (since [90] does not take users' SNR constraint into consideration). Fig. 2-14 and Fig. 2-15 confirm that our design SAG is not only a feasible but also an energy efficient relay deployment strategy for hot-trended wireless relay networks.

Fig. 2-16(a) to Fig. 2-16(d) show the impact of the number of BSs on the network energy consumption. In a $3km \times 3km$ field (Fig. 2-16(a) and Fig. 2-16(b)), we increase the number of BSs from 1 to 10 and observe that, with more BSs deployed in the field, less number of RSs are needed relaying traffic while guaranteeing the SNR for each subscriber. Furthermore, the total power consumption of the placed RSs is decreased as well with more BSs. From Fig.

2-16(b), we can see that more BSs lead to more power cost savings. This is consistent with practical deployment. An area with dense BS deployment needs less RSs to relay the traffic since more traffic can go to BSs directly. We place *relay RSs* on the edges of a minimum spanning tree, the less RSs to place, the less power cost savings could be achieved against the *baseline*. For $5km \times 5km$ field (Fig. 2-16(c) and Fig. 2-16(d)), we increase the number of BSs from 2 to 20, similar results are observed. Fig. 2-16 also confirms that our SAG design outperforms the baseline.

2.2 Power-Aware Microcell BS Placement

2.2.1 Overview

Saving power on BSs becomes a critical issue in wireless cellular networks. Many existing work has proposed to schedule BS into sleep to save energy. However, in reality, it is very difficult to shut down and reboot BSs frequently due to numerous technical issues and performance requirements. In this work, we propose a much more practical solution and offer a new perspective on implementing green wireless networking by embracing the hot-trended small cell network idea. Instead of putting BSs into sleep, we tactically reduce the coverage (and the power usage) of each BS, and strategically place microcells (relay stations) to offload the traffic transmitted to/from BSs in order to save total power consumption.

Small cell network has gain momentum in the past few years and become a hot trend for next generation wireless networking. The authors in [68] tackle the problem of placing the minimum number of relays to achieve bandwidth sufficiency when real-time multimedia streams need to be sent to the sink. They considered the constraints of heterogeneous link capacity and transmission range. Zhang *et al.* in [90] studied a distance-aware relay placement problem in WiMAX Mesh Networks with the goal of placing the minimum number of RSs to cover all the subscribers while satisfying subscribers' data rate request constraint. They considered a two-tiered network and proposed several efficient algorithms to solve the optimization problems on both tiers, respectively. [35] investigated a joint optimization problem on relay node placement and route assignment for two-tiered wireless networks and proposed a recursive weighted clustering binary integer programming algorithm to maximize the total number of information packets received at the BS during the network lifetime. In [26], Elgendy, O.A. *et al.* proposed an optimization framework to maximize either the total cell capacity or the total cell-edge capacity, while taking into consideration the effect of co-channel interference. In [47], the authors investigated the problem of optimal relay placement for coverage extension in relay assisted LTE-A networks. They studied both DL and UL transmission scenarios for optimal relay placement taking into account the SINR of the received signal on the evolved-NodeB (eNB)-RN and RN-User Equipment (UE) links. However, our work is different from all these previous ones. We know that more than 60% power consumption on wireless infrastructure is spent on wireless radio access network and transmission power of BSs is the major part of total power consumption on radio access network. Since transmit power is exponentially increasing with the transmission range, we consider to reduce the transmission range of macro BSs meanwhile deploying multiple micro BSs to support the coverage and users' QoS requirements, which can save power consumption on wireless access network to a great extent. Hence, our purpose of placing relays is to save power consumption in the network, Or we can say, to make BS feel relax while covering all the users each of which has a certain amount of throughputs. We are the first one to propose the new perspective of embracing hot-trended small cells on Green Wireless Networks. However, most of the studies on small cell have focused on network capacity improvement. To the best of our knowledge, nobody has considered how (or if) small cell networks can provide a more energy-efficient greener wireless network.

2.2.2 Problem Statements

In this work, we consider a wireless network with $N(\geq 1)$ BSs and M users. We assume that the locations of BSs and users are known. In other words, users in this work are large static wireless service subscribers, such as Wal-mart, McDonald's and gas stations, which are static but usually have large traffic demands. Since all the users in this work share similar communication characteristics and QoS, we assume that all users have the same data rate service L. Most previous works on green wireless network introduce a sleep mode for BSs. If a BS has a low traffic load or idle during a certain time period, then it could be shut down in order to save wasting power consumption on some components such as air cooling, power amplifier, digital data processing and so on. However, in reality, it is very difficult to shut down and reboot BSs frequently due to numerous technical issues and performance guarantees. In this work, we do not propose to shut down BSs, instead, we "relax" the BSs by reducing their burden of service coverage. Meanwhile, hot-trending small cells (called relay station (RS) in this work) are applied to provide enough coverage with less total network energy consumption. Before we present our problem definitions, let us first introduce the power consumption model for BSs and RSs.

2.2.2.1 Power Models of Base Station and Relay Station

Power consumption of a BS or RS consists of various power costs, including transceiver, power amplifier, digital signal processing, air cooling and so on. Also, it is not only relevant to the transmit power of the antennas but also the traffic loads from the users and some other factors. Many previous literatures [7, 19, 20, 59] were working on it and proposing several useful ones. In our work, we select the power models in [20], [69]. The power model of BS proposed in [20] is:

$$P_{el/macro} = n_{sector}(P_{el/rect} + F(n_{Tx}(P_{el/amp} + P_{el/trans}) + P_{el/proc}) + P_{el/link} + P_{el/airco}$$

with n_{sector} the number of sectors, F the load factor, n_{Tx} the number of transmitting antennas, and $P_{el/rect}$, $P_{el/amp}$, $P_{el/trans}$, $P_{el/proc}$, $P_{el/link}$ and $P_{el/airco}$ the power consumption (in Watt) of the rectifier, the power amplifier, the transceiver, the digital signal processing, the microwave link, and the air conditioning, respectively.

The power model of RS proposed in [20] is:

$$P_{el/micro} = P_{el/rect} + P_{el/airco} + F(P_{el/amp} + P_{el/trans} + P_{el/proc}) \tag{2.1}$$

with F the load factor, and $P_{el/rect}$, $P_{el/airco}$, $P_{el/amp}$, $P_{el/trans}$, and $P_{el/proc}$ the power consumption of the rectifier, the air conditioning, the power amplifier, the transceiver, and the digital signal processing (in Watt), respectively.

Simply put, some energy consumption is related with the transmission (distance and traffic), some are related with the traffic, such as rectifier cost, and others are fixed cost, such as air conditioning cost. Therefore, we simplified our energy consumption model as following:

$$P_{el/macro} = (a_0 r_{bs}^2 + b_0) \times L_{bs} + c_0; \tag{2.2}$$

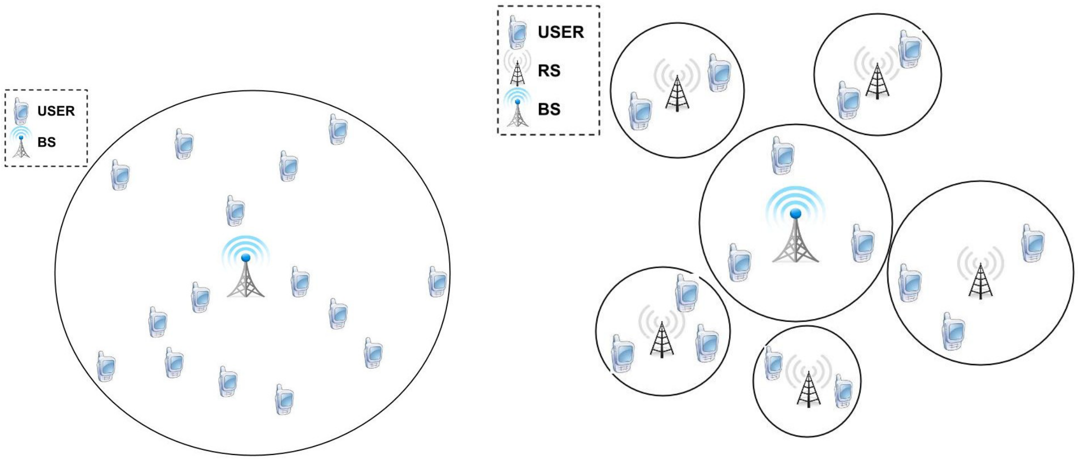
$$P_{el/micro} = (a_1 r_{rs}^2 + b_1) \times L_{rs} + c_1; \tag{2.3}$$

where a_0,a_1,b_0,b_1,c_0,c_1 are constants we can know, r_{bs} and r_{rs} are the transmit range of BS and RS, L_{bs} and L_{rs} are the traffic loads of BS and RS, respectively.

Our power models take transmit range, users' traffic loads, and some constant power costs into consideration. Thus, our models are feasible and practical ones. Furthermore, our power model of BS is consistent with that proposed in [69].

2.2.2.2 Green Relaxed Energy Aware Network Problem

In this section, we present our **G**reen **R**elaxed **E**nergy **A**ware **N**etwork (GREAN) problem. Our goal of our proposed scheme is to tactically reduce the coverage, and then the power consumption, of the BS. For the uncovered users, we strategically place RSs, which have much smaller energy consumption compared to BS, while keeping the total energy consumption reduced. An illustration of our strategy is shown in Fig. 2-17. The traditional BS-takes-all network in Fig. 2-17(a) will be replaced by a hybrid Macro+small cells network in Fig. 2-17(b). Also, in our study, an RS has a maximum transmission range d_R^{max}, but can select its own transmit power range based on coverage designs. Our goal is to minimize the total network energy by strategically placing RSs, as well as adjusting radius (and power usage) of BS and RSs.



(a) BS covers all the users (b) Both BS and RSs cover all the users

Figure 2-17: Illustration of GREAN design

Definition 6 (Green Relaxed Energy Aware Network (GREAN) problem)**.** Given a network with a BS, a set of users $V = \{v_1, v_2, ..., v_n\}$, with traffic rate L, the GREAN problem seeks a network design with K RSs such that:

1. the placement of the K RS, and the transmit power of each placed RS
2. the transmit power of BS
3. User must be covered by BS or RS
4. the total power consumption of BSs and RSs should be minimized. □

It is straightforward to see that the GREAN problem is closely related with K-*center* problem or *Dominant set* problem. Therefore, we speculate that GREAN is also *NP-hard*, and try to present approximation schemes for the problem.

2.2.2.3 A Special Case: GREAG Problem

In the GREAN problem, there are infinite number of locations for placing RSs. To find a final solution for GREAN, first, we try to tackle a problem with limited number of RSs locations, which is defined in following:

Definition 7 (Green Relaxed Energy Aware Grid (GREAG) problem)**.** Given a grid network with grid size d_s, a BS is in the center of the network, and a set of users $V = \{v_1, v_2, ..., v_n\}$, with traffic rate L, locate *on the grids*, shown in Fig. 2-18. The maximum power range of a RS is assumed to be $d_R^{max} = 2 \cdot d_s$, the GREAG problem seeks placement of K RSs and transmit power allocation strategy for RSs and BS such that:

1. *RSs can only be placed in the center of a grid*

2. the placement of K RSs, and the transmit power of each placed RS

3. the transmit power of BS

4. the total power consumption of both BSs and RSs should be minimized. □

Given the maximum coverage range of RSs, d_R^{max}, we can construct a grid network with grid size $d_s = d_R^{max}/2$, as shown in Fig. 2-18. By enforcing a grid network, now we have limited number of potential locations for RSs, which are center of each grid in Fig. 2-18.

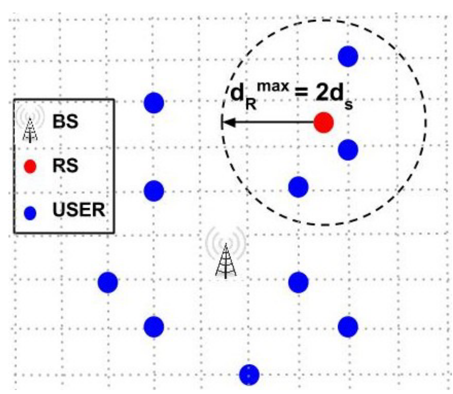


Figure 2-18: An illustration for GREAG problem

For this special case, we first present and prove an approximation solution, which is listed in **Algorithm** 11.

Theorem 2. Algorithm 11 is a $(8\mathcal{M})$-approximation for the GREAG problem. In other words, let the power consumption returned by Algorithm 11 be P, and OPT be an optimal

Algorithm 11: GREAG(d_R^{max}, K)

Let $E(v_i, v_j) \triangleq$ power consumption of RS or BS placed at v_i to cover user v_j;
$V = \{v_0, v_1, ..., v_n\}$ with v_0 is BS, users = $\{v_0, v_1, ..., v_n\}$;
Let $s_0 = v_0$, and $B_0 \leftarrow \{v_1, ..., v_n\}$, the users that covered by B_0;
for $k \leftarrow 0$ to $K-1$ **do**
  $h \leftarrow max\{E(s_j, v_i) \mid v_i \in B_j$ and $0 \leq j \leq k\}$;
  Let v_i be a user whose covering station s_j's energy consumption $E(s_j, v_i)$ would be h;
  v_i needs to be covered by a new RS s_{k+1}; Move v_i to B_{k+1};
  Find the nearest candidate location (center of a grid) to v_i, and place s_{k+1} in the location;
  for each $v_t \in (B_0 \bigcup ... \bigcup B_k)$ **do**
    let j be such that $v_t \in B_j$;
    if $E(v_t, s_j) \geq E(v_t, s_{k+1})$ **and** $dis(v_t, s_{k+1}) \leq d_R^{max}$ **then**
      move v_t from B_j to B_{k+1};
    end if
  end for
end for
if B_0 is empty **then**
  $h_0 \leftarrow min\{E(B_0, v_i) \mid v_i \in B_0 \bigcup ... \bigcup B_{K+1}\}$;
  let v_i be the user who consumes B_0 h_0 and whose energy consumption for its corresponding RS s_l is maximal;
  move v_i from B_l to B_0;
end if
Return $B_0, ..., B_K$;

solution for GREAG, we know $P \leq 8\mathcal{M} \cdot OPT$, where $\mathcal{M} = \max\{\frac{a_0}{a_1}, \frac{b_0}{b_1}, \frac{c_0}{c_1}\}$ from the power model in (2.2)(2.3). □

Proof: For any user i, let OPT_i denote the power consumption for a station (BS or RS) to cover user i in an optimal solution OPT, and let OBJ_i denote the power consumption for a station (BS or RS) to cover user i in our GREAG solution. For a user i in V, let d_{min} denote the distance from i to its covering RS/BS s_i. Following the energy consumption model, OBJ_i, the energy consumption for s_i to cover user i, is

$$OBJ_i \leq (a_0 d_{min}^2 + b_0)L + c_0 \tag{2.4}$$

$$\leq (a_0 (d_R^{max})^2 + b_0)L + c_0 \tag{2.5}$$

$$\leq \mathcal{M} \cdot a_1 (d_R^{max})^2 \cdot L + \mathcal{M} \cdot b_1 \cdot L + \mathcal{M} \cdot c_1 \tag{2.6}$$

Meanwhile, given the fact that users are all on the grids, and RS/BS are in centers of grids, the minimum distance between an RS/BS to a user is $\frac{\sqrt{2}}{2} d_s$, where d_s is the grid size.

Therefore, we have

$$OPT_i \geq (a_1(\frac{\sqrt{2}}{2} \cdot d_s)^2 + b_1)L + c_1 \tag{2.7}$$

According to our assumption, $d_R^{max} = 2 \cdot d_s$, we have

$$\begin{aligned} OPT_i &\geq (a_1(\frac{\sqrt{2}d_R^{max}}{4})^2 + b_1)L + c_1 && (2.8)\\ &\geq (a_1\frac{(d_R^{max})^2}{8} + \frac{b_1}{8})L + \frac{c_1}{8} && (2.9)\\ &= \frac{1}{8}(a_1(d_R^{max})^2 L + b_1 L + c_1) && (2.10) \end{aligned}$$

Combining (2.6) and (2.10), we know that

$$OBJ_i \leq 8 \cdot \mathcal{M} \cdot OPT_i \tag{2.11}$$

Since $P = \sum_{i=1}^{N} OBJ_i$ and $OPT = \sum_{i=1}^{N} OPT_i$, summing all the users, we can see that

$$P \leq 8 \cdot \mathcal{M} \cdot OPT$$

■

2.2.2.4 Solution to GREAN problem

Now in this section we propose a solution to the GREAN problem, where users and BS can be at any place, and there is no constraint on the locations of RSs. Our solution is very similar to the solution to GREAG problem. But the major difference of is that we now *choose the users locations as potential RS placement locations* (shown in Line 8 of Algorithm 12).

Let us use an example in Figs. 2-19 and 2-20 to illustration how the algorithm works. There are six users, $u1 \ldots u6$ in the cell. First, we find the maximum energy for BS to cover a user, which is user u_1 with the longest distance, shown in figure 2-19(a). Then we place an RS, r_1, co-located with u_1. Next, we adjust users coverage comparing energy consumption for each user's from r_1 with energy from BS. Since distances between r_1 and u_4, $u5$ and $u6$ are all larger than RS's maximum transmit range d_R^{max}, they can only be covered by BS. On the other hand, energy consumption of covering u_2 from r_1, $E(r_1, u_2)$, is smaller than

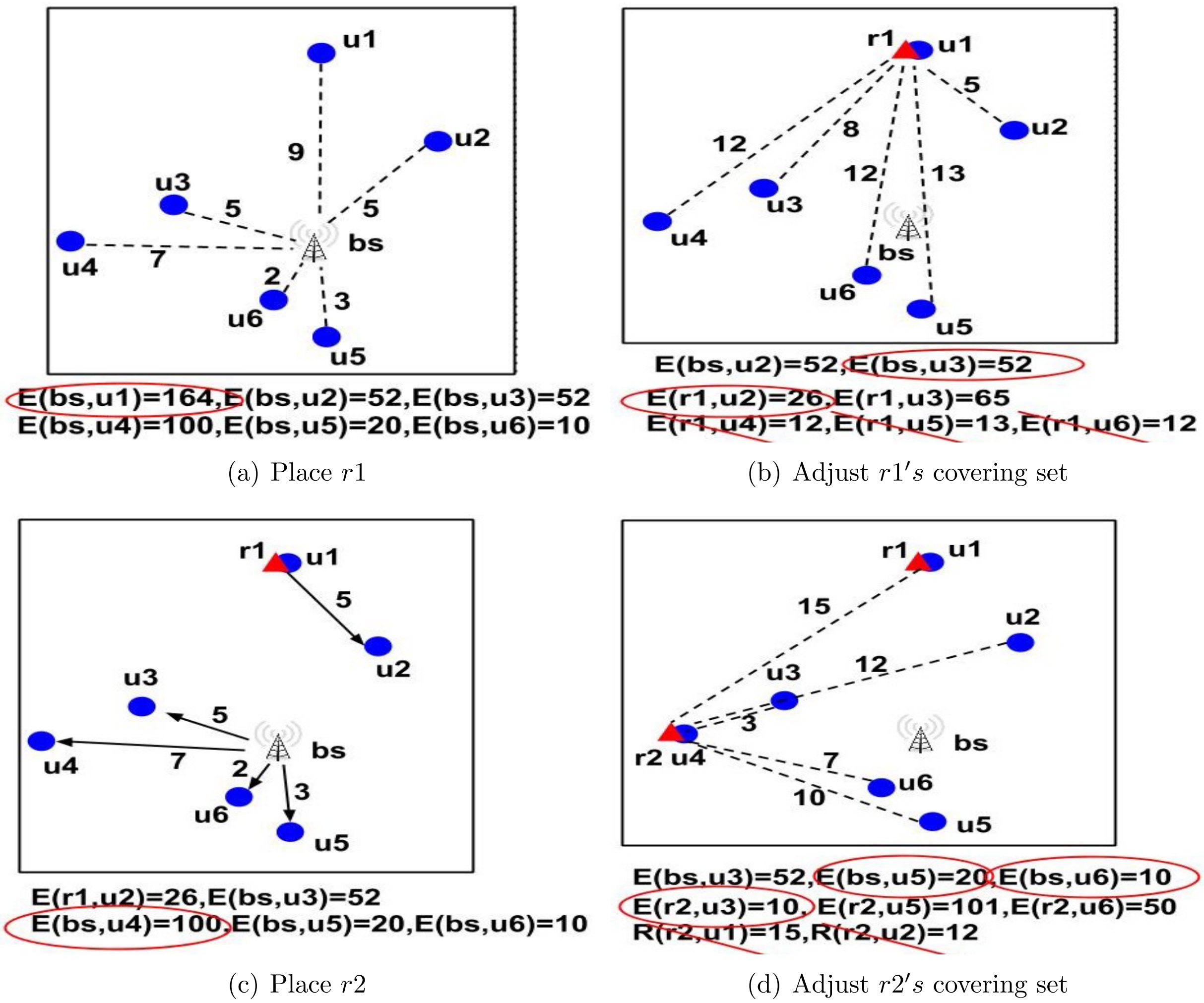
Algorithm 12: GREAN(d_R^{max}, K)

Let $E(v_i, v_j) \triangleq$ power consumption of RS or BS placed at v_i to cover user v_j;
$V = \{v_0, v_1, ..., v_n\}$ with v_0 is BS, users $= \{v_0, v_1, ..., v_n\}$;
Let $s_0 = v_0$, and $B_0 \leftarrow \{v_1, ..., v_n\}$, the users that covered by B_0;
for $k \leftarrow 0$ to $K-1$ **do**
  $h \leftarrow max\{E(s_j, v_i) \mid v_i \in B_j \text{ and } 0 \leq j \leq k\}$;
  Let v_i be a user whose covering station s_j's energy consumption $E(s_j, v_i)$ would be h;
  v_i needs to be covered by a new RS s_{k+1}; Move v_i to B_{k+1};
  Place s_{k+1} at the same location with user v_i;
  for each $v_t \in (B_0 \bigcup ... \bigcup B_k)$ **do**
    let j be such that $v_t \in B_j$;
    if $E(v_t, s_j) \geq E(v_t, s_{k+1})$ **and** $dis(v_t, s_{k+1}) \leq d_R^{max}$ **then**
      move v_t from B_j to B_{k+1};
    end if
  end for
end for
if B_0 is empty **then**
  $h_0 \leftarrow min\{E(B_0, v_i) \mid v_i \in B_0 \bigcup ... \bigcup B_{K+1}\}$;
  let v_i be the user who consumes B_0 h_0 and whose energy consumption for its corresponding RS s_l is maximal;
  move v_i from B_l to B_0;
end if
Return $B_0, ..., B_K$;

$E(bs, u_2)$, the energy consumption of covering u_2 by BS. Therefore, u_2 would be covered by r_1 now. Next, we keep finding the largest energy consumption for a RS/BS to cover any of its covering user. Now $E(bs, u_4)$ has the maximum energy. Then we place another RS r_2 at u_4's location, shown in Fig. 2-19(c). Now we have r_1 covers u_1, u_2, r_2 covers u_3, u_4 and BS covers u_5, u_6. Following same process, we place r_3 to cover u_2, r_4 to cover u_5, u_6, and leave no user to be covered by BS. In order to keep BS alive, we force BS to cover one nearest user with the minimum energy. In Fig. 2-20(d), BS covers u_6 since $u6$ is the one nearest to BS. Final solution is shown in Fig. 2-21.

Theorem 3. Algorithm 12 is a $(1+\alpha)$-approximation for the GREAN problem. More specifically, if the power consumption of both BS and RSs returned by Algorithm 12 is denoted as P, we have $P \leq (1+\alpha) \cdot OPT$, where $\alpha = \frac{a_1 R_{max}^2 L}{b_1 \cdot L + c_1}$ and OPT is an optimal solution for the GREAN problem. □

Proof: For any user i, we let OPT_i denote the power consumption for a station (BS or RS) to



(a) Place $r1$ (b) Adjust $r1's$ covering set

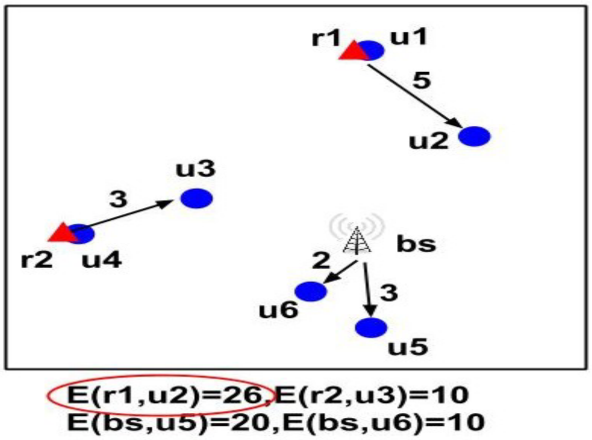
(c) Place $r2$ (d) Adjust $r2's$ covering set

Figure 2-19: Illustration of GREAN (place $r1$ and $r2$)

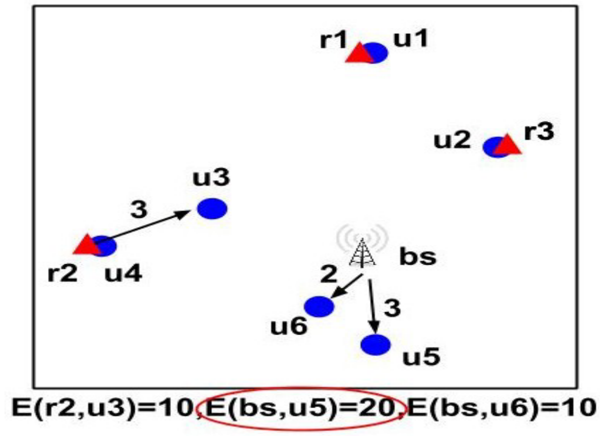
cover user i in the optimal solution, and OBJ_i denote the power consumption for a station (BS or RS) to cover user i in our GREAN solution. Since we know that $a_0 \geq a_1, b_0 \geq b_1, c_0 \geq c_1$, if user i is covered by BS in an optimal solution, we have

$$
\begin{aligned}
OPT_i &= (a_0(r_{bs}^{opt})^2 + b_0)L + c_0 \qquad (2.12)\\
&\geq b_0L + c_0 \geq b_1L + c_1
\end{aligned}
$$

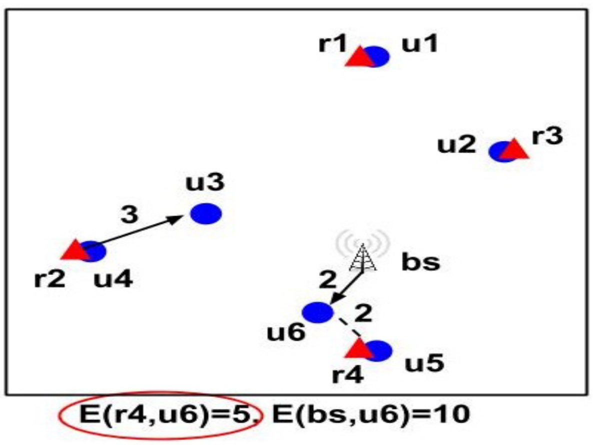
where r_{bs}^{opt} is the distance between BS and i in the optimal solution.



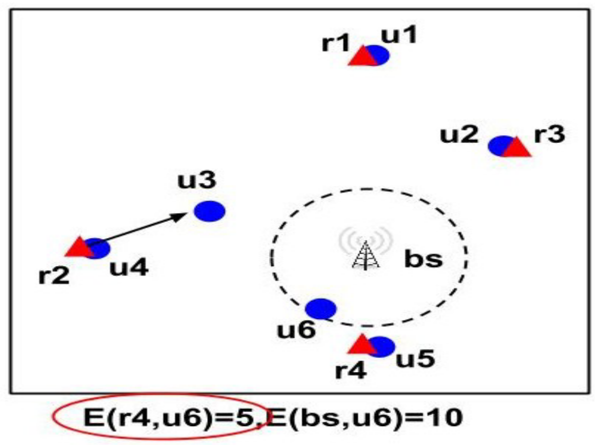
(a) Place $r3$



(b) Adjust $r3's$ covering set



(c) Place $r4$



(d) Adjust $r4's$ covering set

Figure 2-20: Illustration of GREAN (place $r3$ and $r4$)

If user i is covered by a RS in optimal solution, we also have

$$
\begin{aligned}
OPT_i &= (a_1(r_{rs}^{opt})^2 + b_1)L + c_1 \\
&\geq b_1 L + c_1
\end{aligned}
\tag{2.13}
$$

where r_{rs}^{opt} is the distance between an RS and i in the optimal solution. Thus, no matter if user i is covered by BS or RS, we have

$$
OPT_i \geq b_1 L + c_1 \tag{2.14}
$$

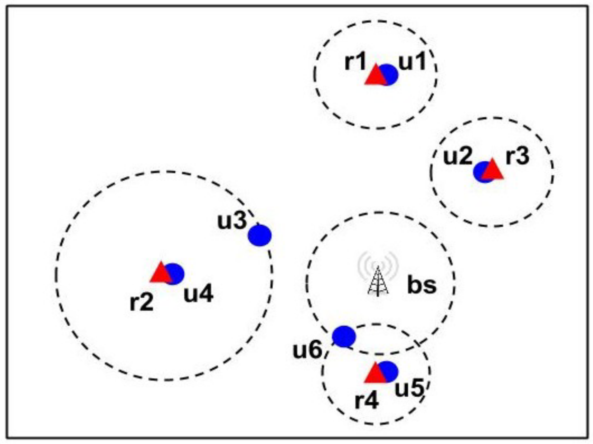


Figure 2-21: Illustration of covering sets

Case 1: If in our GREAN solution, user i is covered by an RS. Then we have

$$OBJ_i \leq (a_1 R_{max}^2 + b_1)L + c_1$$

Therefore, in this case, we have

$$\frac{OBJ_i}{OPT_i} \leq \frac{(a_1 R_{max}^2 + b_1)L + c_1}{b_1 L + c_1} = (1 + \frac{a_1 R_{max}^2 L}{b_1 \cdot L + c_1})$$

Let $\alpha = \frac{a_1 R_{max}^2 L}{b_1 \cdot L + c_1}$, we have

$$OBJ_i \leq (1 + \alpha) OPT_i$$

Case 2: If in our GREAN solution, user i is covered by BS. Let d_{min} be the distance between i and its nearest RS placed in GREAN solution. Following our algorithm, the reason for a user i covered by BS is because

$$\begin{aligned} OBJ_i = (a_0 \cdot d_{BS,i}^2 + b_0)L + c_0 &\leq (a_1 d_{min}^2 + b_1)L + c_1 \\ &\leq (a_1 R_{max}^2 + b_1)L + c_1 \end{aligned} \tag{2.15}$$

Like in Case 1, $OBJ_i \leq (1 + \alpha) OPT_i$ in Case 2.

Therefore, for each user i, we have $OBJ_i \leq (1 + \alpha) OPT_i$. Since $P = \sum_{i=1}^{N} OBJ_i$ and $OPT = \sum_{i=1}^{N} OPT_i$, summing all the users, we can easily see

$$P \leq (1 + \alpha) OPT$$

where $\alpha = \frac{a_1 R_{max}^2 L}{b_1 \cdot L + c_1}$. ■

2.2.2.5 Minimize the number of Relay Stations for GREAN

In previous sections, we have studied the case placing a fixed number,K, of RSs in the network. In this section, we try to see if we can use less number of RSs ($\leq K$) for better solutions.

Definition 8 (Budget Aware Power Saving (BAPS) problem)**.** Given a network with a BS, a set of users $V = \{v_1, v_2, ..., v_n\}$ with traffic rate L the BAPS problem seeks an optimal amount of relay stations R and transmit power allocation strategy for R such that:

1. the placement of $\leq K$ RSs
2. the transmit power of each placed RS
3. the transmit power of BS
4. the total power consumption of both BSs and RSs should be minimized. □

Based upon the observation that we obtain the results from GREAN problem (using a fixed K RSs), we present a solution to the BAPS problem based upon GREAN solution via a linear search on the number of RSs to place.

Algorithm 13: Budget Aware Power Saving (BAPS)(d_R^{max}, K)

$P \leftarrow \emptyset$;
for $k \leftarrow 1$ to K **do**
    $p_k \leftarrow GREAN(d_R^{max}, K)$;
    $P \leftarrow P \bigcup p_k$;
end for
$P_{min} \leftarrow$ the minimum value in P;
return P_{min};

2.2.2.6 Multi-cell Scenarios

So far, we have studied all the problems in a single cell. But in practice, network carriers will deploy multiple macro cells in a market. The number of RS will be used for the whole market (multiple cells), instead of for each cell. Therefore, in this section, we study how to set up RSs for the multi-cell scenario.

Definition 9 (Multi-Cell Budget Aware Power Saving (MC-BAPS) problem)**.** Given B BSs in a network market, a set of users covered by these BSs, a total number of RSs (or micro cells) that can be distributed in the market (among B BSs). the MC-BAPS problem seeks:

1. the amount of RS (micro cells) to be allocated for each BS (macro cell)

2. the placement of RSs for each BS (macro cell)

3. the transmit power of each placed RS in each BS (macro cell)

4. the transmit power of each BS

5. the total power consumption of all BSs and RSs should be minimized □

In order to solve the MC-BPAS problem, one intuitive solution is to convert the multi-cell problem into the previously studied single cell problem. If the number of placed RS for each single cell is known, we can use BAPS algorithm to solve the single cell power saving problem. One major question to answer is *how to allocate RSs among multiple BSs (macro cells)?*. In other words, we need to know how many RSs (micro cells) to be used in each BS (macro cell), with the constraint that the total RSs placed cross the whole market is no more than K. We start by distributing RSs evenly among all of the B BSs. Then, we will try to check if we can achieve more savings if we move an RS from one macro cell (BS) to another. We will keep redistributing RSs among macro cells to save power until no more power can be saved.

Let us use an example in Fig. 2-22 to illustrate our algorithm. There are 4 cells in this market, denoted sa c_1, c_2, c_3, c_4. The maximum number of RSs K in this market is 16. Initially, we equally divide 16 RSs into each cell. So each cell has 4 RSs at the beginning. Then, for each cell we calculate the power consumption with 3 RSs, 4 RSs and 5 RSs. Next, we check if redistribution of RS numbers will provide more power savings. If C_2 reduce its RS to 3, while C_1 gets one more RS, we can see $P|_{k1=5} = 5$ and $P|_{k2=3} = 2$. Since $[P|_{k1=5}+P|_{k2=3}] = 7 > [P|_{k1=4}+P|_{k2=4}]$, no power can be saved during this RS redistribution.

Algorithm 14: MC-BAPS(d_R^{max}, K)

Initialize S_i and k_i to be empty;
for each user v_i in U **do**
  Assume user v_i access to the nearest BS bs_j: $S_j \leftarrow S_j \bigcup v_i$;
end for
Evenly distribute $\lfloor \frac{K_{budget}}{|L_{bs}|} \rfloor$ RSs into each BS;
while not DONE **do**
  $P \leftarrow \emptyset$;
  for each BS bs_i **do**
    $p_i \leftarrow \mathsf{BAPS}(d_R^{max}, k_i)$;
    $p'_i \leftarrow \mathsf{BAPS}(d_R^{max}, k_i + 1)$;
    $p''_i \leftarrow \mathsf{BAPS}(d_R^{max}, k_i - 1)$;
  end for
  for each BS bs_i **do**
    for each BS bs_j $(j \neq i)$ **do**
      if $(p'_i + p''_j) < (p_i + p_j)$ **then**
        Record energy saving $\triangle P_{ij} = [(p_i + p_j) - (p'_i + p''_j)]$ and add it into P;
      end if
    end for
  end for
  if P is empty set **then**
    DONE == 1;
  else
    pick $\triangle P = \max\{P_{ij} \forall i, j\}$;
    //Reduce 1 RS from bs_j, and shift the quota (1 more RS) to bs_i;
    $k_i \leftarrow k_i + 1$; $k_j \leftarrow k_j - 1$;
  end if
end while
for each BS bs_i **do**
  $p_i \leftarrow BAPS(d_R^{max}, k_i)$;
end for
$P_{total} \leftarrow 0$;
for each BS bs_i **do**
  $P_{total} \leftarrow P_{total} + p_i$;
end for
return P_{total};

Thus, no RS redistribution between these two cells. However, RS redistribution between some cells do provide power savings, such as $\{3 \mid C_3 \rightarrow C_1\}$, $\{5 \mid C_3 \rightarrow C_2\}$ and $\{5 \mid C_3 \rightarrow C_4\}$. After RS redistribution, the number of RS in each cell is determined. Then, we will apply BAPS algorithm for each cell to find RS placement and power consumptions.

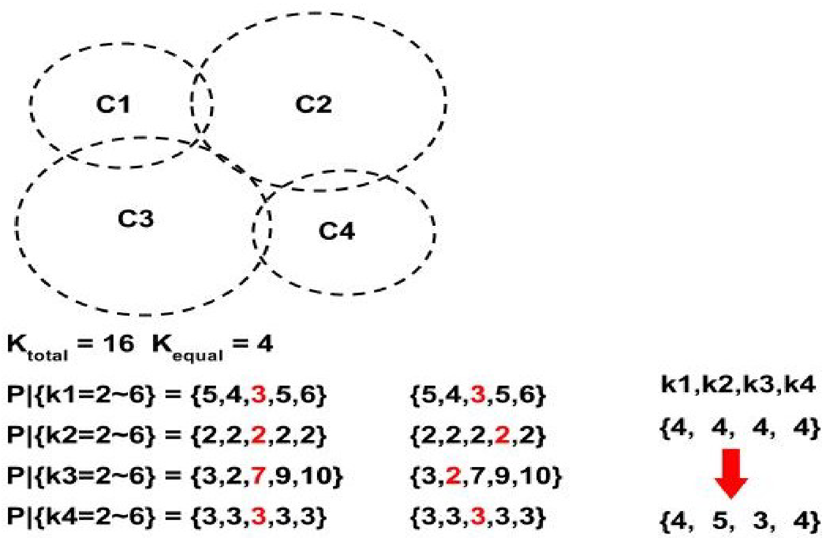


Figure 2-22: Illustration of MC-BAPS

2.2.3 Numerical Results

In this section, numerical results are presented to show the effectiveness of our schemes, including GREAN, BAPS, MC-BAPS algorithms. All the simulations are run on a Intel Core i7 CPU of 2.7GHz with 8GB types of memory. All the users and BSs are uniformly distributed in a square testing field. All the figures illustrate the average of 10 test runs for various scenarios.

2.2.3.1 Simulation Environment Settings

In order to test the performance of GREAN and BAPS algorithms, we set a square region with size of $2km \times 2km$. The BS is placed at the centre of this region. We assume that this BS covers all the users in this region. There are $[50 - 550]$ users uniformly distributed in this two-dimensional region. Each user has a throughput of $[0.4Mbps - 1.4Mbps]$. The number of RSs K is changing from 3 to 36 for testing GREAN in this single cell. The maximum transmit range of RS is set to $300m$. In testing MC-BAPS algorithm, we set a multi-cell region with size of $3km \times 3km$. 4 BSs are uniformly distributed in this two-dimensional region. The number of users in multi-cell region is changing from 200 to 2200. The maximum number of RSs in this multi-cell region is changing from 20 to 130. We refer to [20], [19] and [69] for the parameters of our power models. We list all the parameters which are used in our power models in Table 2.3.

Symbols	Macrocell	Microcell
n_{sector}	3	n/a
n_{Tx}	2	n/a
$P_{el/rect}$	100 W	100 W
$P_{el/amp}^{max}$	156.3 W	16.6 W
$P_{el/trans}$	100 W	100 W
$P_{el/proc}$	100 W	100 W
$P_{el/link}$	80 W	n/a
$P_{el/airco}$	1500 W	60 W
R_{max}	1000 m	300 m
a	1.95e-6	7.7e-7
b	1.875	0.8
c	605	60

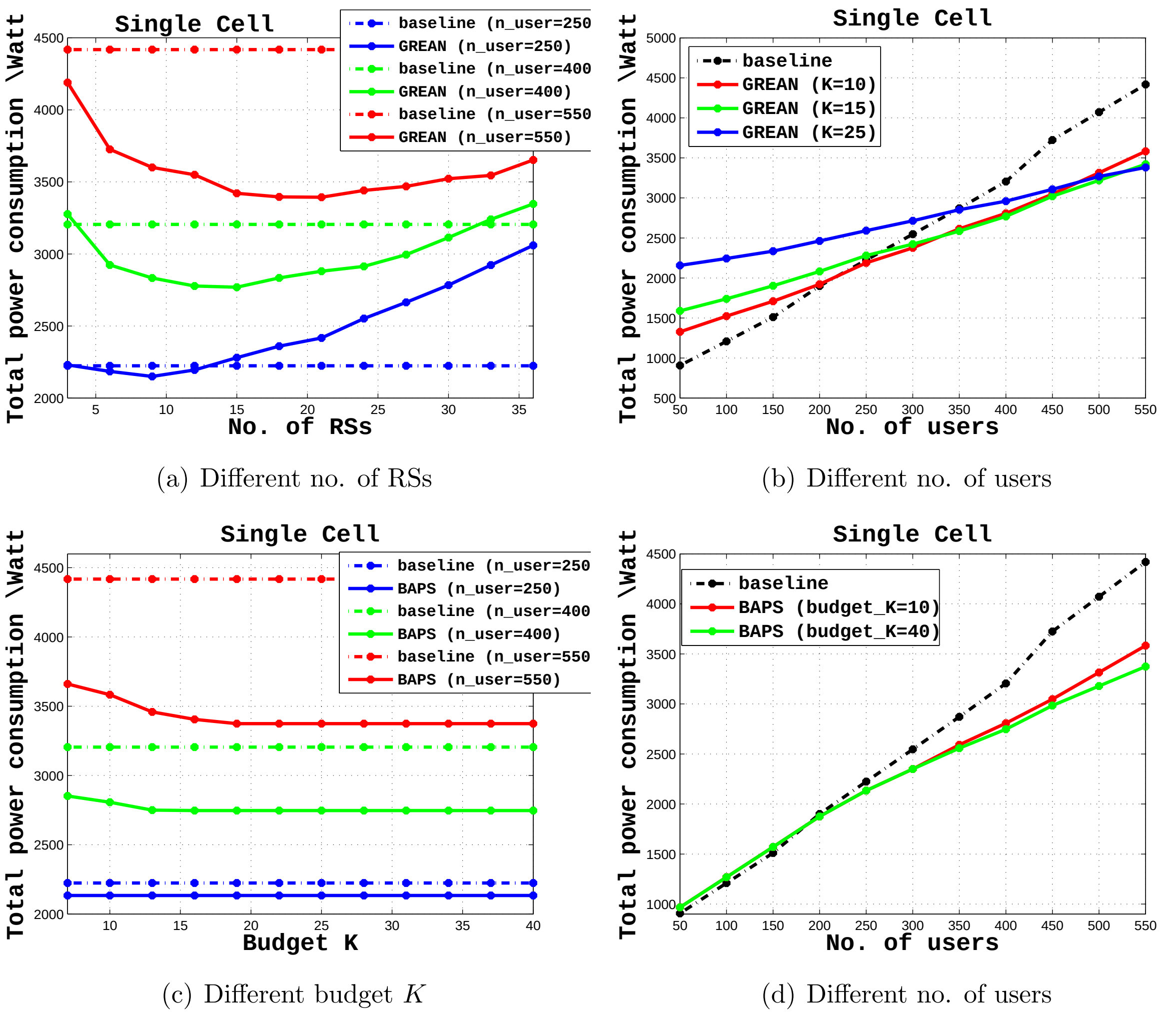
Table 2.3: Power model parameters

In our work, we are targeting 4G LTE macrocell BS with 2×2 MIMO antennas in each sector. There are total three sectors in each BS. We present the evaluations of our schemes in the following sections.

2.2.3.2 Evaluation of GREAN algorithm

From Fig. 2-23(a) and 2-23(b), we can see that GREAN algorithm has good performance on power saving. Comparing with baseline in which no RS exists and all the users must get access to BS, GREAN can save up to 23% power consumption. More power can be saved especially when there are more users in the cell since power saving from more users can better counterbalance the static power consumption of placing K RSs.

If we are given the exact number of RSs to place, network operator need to pay the static power consumption of these K RSs, air conditioning and rectifier consumption for instance. The dynamic power saving of each user can be seen as the difference between the power consumption for BS to cover this user and that for one RS to cover this user. If sum of the dynamic power saving for all the users is more than all the static power consumptions of K



(a) Different no. of RSs

(b) Different no. of users

(c) Different budget K

(d) Different no. of users

Figure 2-23: GREAN and BAPS in different scenarios

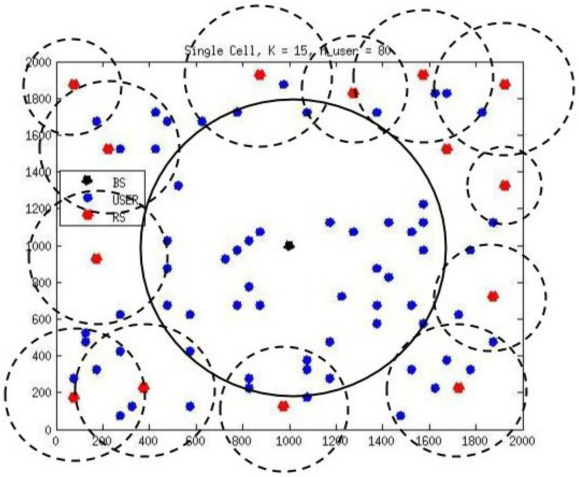
RSs, then we can say that total power consumption from current placements is lower than baseline. Thus, for the case of only a few users in the cell, sum of the dynamic power saving from all the users cannot counterbalance the static power consumptions of K RSs. Total power consumptions of both BS and K RSs would be larger than baseline. In other words, no power can be saved especially for the cell with few users. It can be easily seen from Fig. 2-23(a) and 2-23(b).

In Fig. 2-23(b), we can easily find that when the number of users is less than 250, placing more than 10 RSs will cause over-baseline power consumptions. Our scheme does not work well in these cases. Moreover, larger K would require more users to counterbalance the static power consumptions in order to achieve power saving. It can be seen from Fig. 2-23(b) between blue line and red/green line.

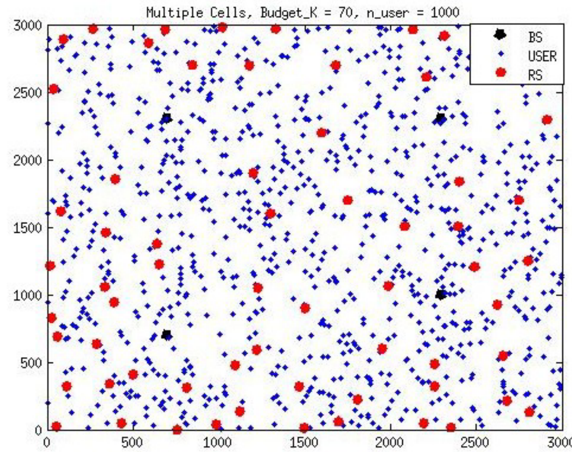
In Fig. 2-23(a), we can see a trend of power consumptions for a fixed number of users with increasing the number of RSs placed. More RSs being placed does not mean more power can be saved in the cell. For a fixed number of users, placing the first few RSs will easily achieve power saving since the dynamic power saving for covering all the users can easily counterbalance the static power consumption of these few RSs. With more RSs placed in this cell, the static power costs will grow linearly with the number of RSs while the dynamic power saving will grow more and more fast at the beginning and then grow slowly down. Thus we can see the trend. From this trend, we can know for a single cell with fixed number of users, there would be a maximum power saving corresponding to a certain number of RSs placed in. We illustrate the GREAN results in Fig. 2-24(a) with the number of RSs fixed to 15 and that of users fixed to 80. We can see the coverage circles of each station in this cell.

2.2.3.3 Evaluation of BAPS algorithm

From Fig. 2-23(a) we know, there would be an optimal number of RSs achieving the most power saving in a fixed user scale. Since BAPS is based on linear searching on the number of RSs and selecting the minimum power consumption as its result, it will always return the



(a) GREAN results



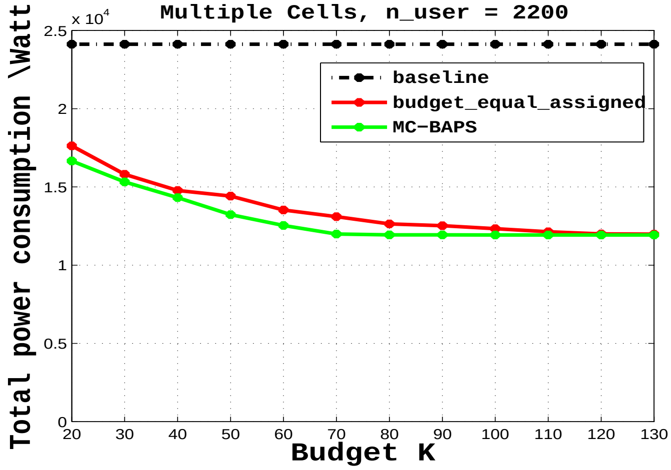
(b) MC-BAPS results

Figure 2-24: Illustration of GREAN and MC-BAPS results

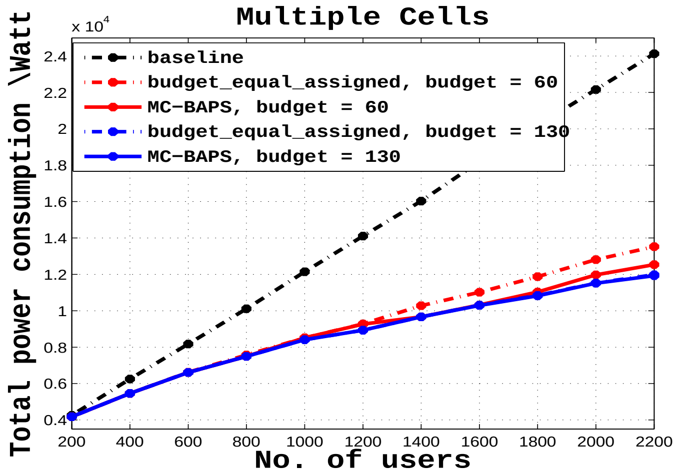
same result if the given maximum number of RSs is more than the optimal number of RSs. Hence we can see in Fig. 2-23(c), the solid red line is decreasing with maximum number of RSs increasing until K_{max} reaches 19. We can say, for this single cell with a user scale of 550, in order to achieve the most power saving, we need to place 19 RSs in this cell. We denote the optimal number of RSs as K_{opt}. From Fig. 2-23(c), we can see $K_{opt} = 19$ for a user scale of 550, $K_{opt} = 13$ for a user scale of 400, and $K_{opt} < 10$ for a small user scale of 250. From Fig. 2-23(d) we see, in a user scale larger than 200, we can see BAPS has good power saving comparing with baseline. The more power saving, the more users get involved in this cell.

2.2.3.4 Evaluation of MC-BAPS algorithm

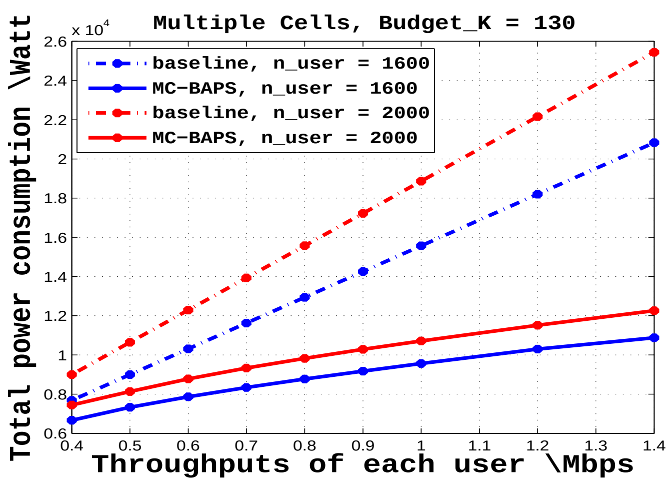
In the multi-cell cases, we add one more baseline in which we set the maximum number of RSs in each cell is equally divided from the maximum number of RSs in this region. Comparing with this new added baseline, we can see how much our MC-BAPS algorithm outperforms the equal assignment solution. From Fig. 2-25(a) and 2-25(b), we see both MC-BAPS and equal assignment solution outperforms baseline pretty much. Furthermore, MC-BAPS outperforms equal assignment solution up to 11%. MC-BAPS saves more power



(a) Different K_{total}



(b) Different no. of users



(c) Different throughputs of each user

Figure 2-25: Results from multiple cells comparing with equal assignment solution and baseline in a larger user scale. From Fig. 2-25(b) we can see, when $K_{total} = 130$, MC-BAPS and equal assignment solution has the same performance since K_{opt} for each cell is much smaller than 32 ($\lfloor \frac{130}{4} \rfloor$). Each cell achieves the same power saving with either one more RS placed in or one less RS placed in so that no K transference is needed. Without K transference, our MC-BAPS makes no improvement against equal assignment solution. They have the same results.

In Fig. 2-25(c), we vary the throughputs of each user from $0.4Mbps$ to $1.4Mbps$. With a larger throughput, both baseline and our scheme will have a higher power consumption since BS or RSs need to consume more power to maintain the communication link with its covering users when they have a larger traffic demand. However, our scheme can achieve a larger power saving when each user has a higher throughput, as can be easily seen in Fig. 2-25(c). Also, in a multi-cell region, we can find that our scheme can have more than 50% power saving in some scenarios. All the numerical results show that our schemes outperforms baseline pretty much. They can achieve a large amount of power saving, especially in a large user scale. We illustrate the MC-BAPS results in Fig. 2-24(b) with maximum number of RSs for this multi-cell region fixed to 70 and the number of users set to 1000. From Fig. 2-24(b), we can see 4 BSs are located near the four corners of this region. All the red points construct a RS set which is our MC-BAPS result.

Chapter 3

Joint Resource Optimization for Green D2D Communications

3.1 Overview

Device-to-Device (D2D) communication commonly refers to the technique that enables wireless devices to communicate directly with each other without an infrastructure of access points or BSs, which has been considered as a key enabling technology for the next generation, (i.e., 5G) wireless communications. We illustrate the concept of D2D communications as an underlay to a cellular network in Fig. 3-1.

Basically, in such a network, two User Equipment (UE) units can communicate directly with each other over the *D2D link*. The BS only helps UE units set up connections without relaying any data traffic. With D2D communications, a UE unit can transmit packets to another UE unit nearby at a reduced power level such that power consumption can be reduced, moreover, interference to other traditional communications (via a BS) can be mitigated, which can improve network capacity. In addition, D2D communications can offload the traffic of BS, which can improve network capacity further.

To take full advantage of D2D communications, channels need to be carefully allocated

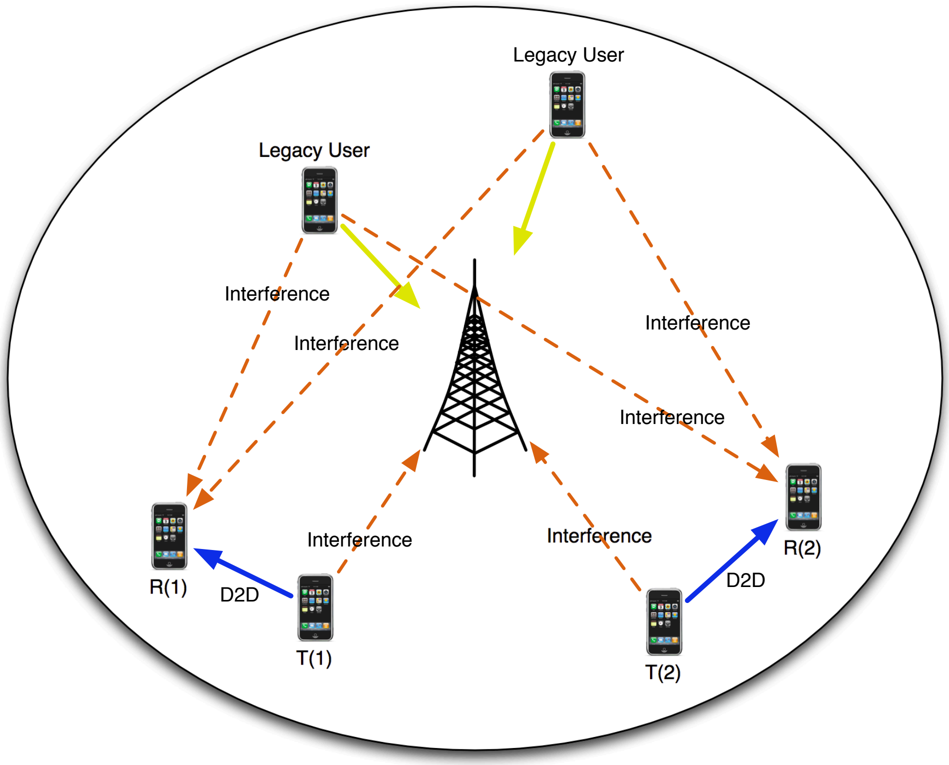


Figure 3-1: A wireless network with D2D links

and transmit power (per channel) needs to be carefully assigned. However, resource allocation in a wireless network with D2D links is different from that in traditional wireless networks because of there exists an additional problem, mode selection (i.e., determining the mode, D2D or cellular, to be used for data transmissions), which is coupled with other resource allocation problems. Even though mode selection has been addressed by a few recent papers [37, 46, 82, 83], some of them [37, 46] were focused on 3G WCDMA-based cellular networks in which data transmissions were conducted on a single channel and the others [82, 83] aimed at maximizing the sum of data rates of mobile users (network throughput). In this work, we study a new optimization problem for D2D communications in an OFDMA-based wireless network, whose objective is to minimize total power consumption while meeting user data rate requirements. We propose a joint optimization approach to solve it, which jointly determines mode selection, channel allocation and power assignment. In addition, the Shannon's equation has been widely adopted to model link data rate, which is not practical since it provides an upper bound, rather than the actual value, for link data rate. We, however, consider a practical model, in which link data rate is an increasing step function of Signal to Interference and Noise Ratio (SINR). Practically, by leveraging the Adaptive Modulation and Coding (AMC) technique, link data rate on a sub-channel becomes a discrete increasing step function $C(\cdot)$ of the SINR (at the receiver). For example, 5 SINR thresholds are specified in the WiMAX standard [43], each of which corresponds to a different modulation index and data rate.

The differences between this work and related papers are summarized as follows: 1) Unlike some related works studying D2D communications in a single-channel 3G CDMA-based cellular network [37,46], we consider an OFDMA-based cellular network with multiple sub-channels and study channel allocation. 2) Unlike some related works that aimed to improve network capacity/throughput [17, 38, 53, 82, 83, 85, 86], the main objective here is to minimize total power consumption to enable green wireless networking. 3) Unlike most related works [17,22,38,46,53,57,82,83,85,86] that modeled link data rate using a continuous function based on Shannon's theorem, we consider a practical model in which link data rate is an increasing step function of SINR. 4) We aim to present a practical algorithm with low time complexity, which is different from the exhaustive search based approach [46] with high time complexity. 5) The problem studied here is mathematically different from those optimization problems formulated based on game theory in [3,92], or the problems studied in [6,91] (which have different objective functions and constraints). 6) In addition, different from our early work [31], we propose a new algorithm, Joint-2, to solve the problem, and present more simulation results for justification based on both the WiMAX and LTE standards in this work.

The rest of this chapter is organized as follows. The system model and problem formulation are presented in Section 3.2. The joint optimization algorithms that we propose are discussed in Section 3.3, which is followed by performance evaluation in Section 3.4.

3.2 System Model and Problem Formulation

Table 3.1: Major Notations

Notation	Description
$C(\cdot)$	The per-channel link data rate function;
$G_{T(i),R(j),k}$	Channel gain between transmitter $T(i)$ of link i and receiver $R(j)$ of link j on sub-channel k
$I^{\text{legacy}}_{R(i),k}$	Interference to receiver of D2D link i on sub-channel k contributed by legacy users;
K	The number of sub-channels;
m_i	The mode of D2D link i;
N	The number of D2D links;
$p_{i,k}$	The transmit power of D2D link i on sub-channel k.
$p_{i,m,k}$	The transmit power of D2D link i on sub-channel k with mode m.
P^{legacy}	The maximal allowable interference power on legacy sub-channel;
Q^{legacy}	The set of sub-channels allocated to legacy links;
$R_{i,0}/R_{i,1}$	The data rate of D2D link i in cellular or D2D mode;
$R(\cdot)$	The receiver of a D2D link;
$T(\cdot)$	The transmitter of a D2D link;
Γ_i	The data rate requirement of D2D link i.

In this work, we consider a single cell in an OFDMA cellular network, which consists of a BS, N pairs of D2D users (a.k.a D2D links), M legacy users (which only communicate with the BS), and K non-overlapping sub-channels. Since neighboring cells can be allocated sets of different channels, they can be operated independently in an interfere-free manner. Each D2D link i consists of a D2D transmitter $T(i)$ and a D2D receiver $R(i)$. Similar to [17,46], we focus on uplink communications since we aim to minimize power consumption of UE units by leveraging D2D communications. Each D2D link i can work in one of the two modes: 1) D2D mode: $T(i)$ directly communicates with $R(i)$; 2) cellular mode: $T(i)$ communicates with $R(i)$ via the BS (as relay). A subset of available sub-channels are assumed to be taken by legacy users for serving their own traffic, which can be re-used by D2D links as long as the total power contributed by D2D links does not exceed a given threshold P^{legacy} at the BS. A relatively conservative threshold can be set to guarantee that traditional communications are not affected by D2D communications.

$G_{T(i),R(i),k}$ denotes the gain of a link i on sub-channel k, which can be measured periodically using pilot signals. $p_{i,k}G_{T(i),R(i),k}$ gives power received at $R(i)$ on sub-channel k, and $p_{j,k}G_{T(j),R(i),k}(j \neq i)$ gives the interference contributed by $T(j)$ at $R(i)$ on sub-channel k, where $p_{i,k}$ is the transmit power at $T(i)$ on sub-channel k. In closely related works [17,38], the well-known Shannon's equation is used to calculate link data rate. However, it is known that the Shannon's equation gives the capacity of a link, which may not be achievable in practice. Moreover, link data rate is usually not a continuous function of the Signal-to-Interference-Plus-Noise-Ratio (SINR). As mentioned above, we consider this practical model for link data rate. So if the SINR and spectrum bandwidth of a sub-channel k of link i are given, then we can obtain the data rate of link i on sub-channel k via the function $C(\mathrm{SINR}_{i,k})$, which can be at several different *levels*. In the simulation, we adopted discrete increasing step functions based on both the WiMAX and LTE standards, which will be described in greater details in Section 3.4. A data rate constraint needs to be enforced for each D2D link i, which requires its data rate to be no less than a given threshold Γ_i.

Here, we aim to solve an optimization problem, which minimizes total transmit power of UEs subject to a data rate requirement for each UE based on the practical link data rate model mentioned above. Solving this problem can enable power-efficient D2D communications with satisfying data rates for mobile users. We present the problem formulation formally in the following, which is referred to as the *Green-D2D* problem.

- Mode selection variables $\mathbf{m} = \{m_i | m_i = \{0, 1\}, 1 \leq i \leq N\}$: $m_i = 1$ if D2D link works in the D2D mode; $m_i = 0$, otherwise.

- Channel-power assignment variables $\mathbf{p} = \{p_{i,k} \geq 0 | 1 \leq i \leq N, 1 \leq k \leq K\}$: $p_{i,k}$ gives $T(i)$'s transmit power on sub-channel k. Note that $p_{i,k} = 0$ if sub-channel k is not allocated to D2D link i.

Green-D2D

$$\min_{\langle \mathbf{m}, \mathbf{p} \rangle} \sum_{i=1}^{N} \sum_{k=1}^{K} p_{i,k} \tag{2.1}$$

Subject to:

$$m_i R_{i,1} + (1 - m_i) R_{i,0} \geq \Gamma_i, \forall i \in \{1, \cdots, N\}; \tag{2.2}$$

$$\sum_{i=1}^{N} m_i p_{i,k} G_{T(i),BS,k} \leq P^{\text{legacy}}, \forall k \in Q^{\text{legacy}}; \tag{2.3}$$

$$(1 - m_i) p_{i,k} \sum_{j \neq i} (1 - m_j) p_{j,k} = 0, \forall i, j \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\} \backslash Q^{\text{legacy}}; \tag{2.4}$$

$$(1 - m_i) p_{i,k} = 0, \forall i \in \{1, \cdots, N\}, \forall k \in Q^{\text{legacy}}; \tag{2.5}$$

$$\sum_{k=1}^{K} p_{i,k} \leq P^{\max}, \forall i \in \{1, \cdots, N\}. \tag{2.6}$$

where:

$$R_{i,0} = \sum_{k \in \{1, \cdots, K\} \backslash Q^{\text{legacy}}} C\left(\frac{p_{i,k} G_{T(i),BS,k}}{\sum_{j \neq i} p_{j,k} G_{T(j),BS,k} + N_0}\right), \forall i \in \{1, \cdots, N\}; \tag{2.7}$$

$$R_{i,1} = \sum_{k \in \{1, \cdots, K\}} C\left(\frac{p_{i,k} G_{T(i),R(i),k}}{\sum_{j \neq i} p_{j,k} G_{T(j),R(i),k} + I^{\text{legacy}}_{R(i),k} + N_0}\right), \forall i \in \{1, \cdots, N\}. \tag{2.8}$$

Note that in the formulation, we use $R_{i,0}$ (Equation (2.7)) to denote the rate of D2D link i working in the cellular mode; and $R_{i,1}$ (Equation (2.8)) to denote the rate of D2D link i working in the D2D mode. In Equation (2.8), $I^{\text{legacy}}_{R(i),k} = 0, k \in \{1, \cdots, K\} \backslash Q^{\text{legacy}}$. The objective (2.1) is to minimize the total power consumption of D2D links. The following constraints must be satisfied:

- *Link data rate constraints (2.2):* The data rate of each D2D link is no less than the given threshold Γ_i. As mentioned above, the per-channel link data rate is given by a discrete increasing step function $C(\cdot)$ of the SINR and sub-channel index.

- *Interference constraints (2.3):* On each sub-channel used by legacy users, the total interference power contributed by all links working in the D2D mode should not exceed the given threshold P^{legacy}.

- *Channel allocation constraints (2.4) and (2.5):* Sub-channels allocated to the legacy links cannot be used for D2D links working in the cellular mode. Moreover, two D2D links both working in the cellular mode can not share a common channel.

- *Power assignment constraints (2.6):* The transmitter of each D2D link distributes its power to the set of assigned sub-channels and the sum of the power assigned to these sub-channels cannot exceed the maximum power level $P^{\max}$.

This problem is a non-linear integer programming problem, which is usually very hard to solve. So we present effective and efficient heuristic algorithms to solve it in polynomial time.

3.3 Joint Optimization Algorithms

The Green-D2D problem can be easily divided into 3 subproblems: *mode selection, channel allocation and power assignment*. A trivial solution is to solve the problem in three separate steps and then combine solutions to the three subproblems together. However, such a method usually does not work well, which has been validated by our simulation results. We present two algorithms, which solve these three subproblems jointly.

3.3.1 Joint Algorithm 1 (Joint-1)

In this algorithm (denoted as *Joint-1*), we use linear search to determine transmission modes (D2D or cellular) using power consumption as guidance first and then jointly compute the channel allocation and power assignment accordingly.

The goal of the mode selection subproblem is to find a solution which can potentially lead to a low-power channel-power assignment. The mode selection is a combinatorial problem. It is not possible to examine all the combinations since the total number of such combinations increases exponentially with the number of D2D links (N). We certainly want a

D2D link to work on a mode with low power consumption. However, it is hard to obtain its power consumption without knowing transmission modes, channel allocations and power assignments of other links. Our idea for mode selection is to sort all the D2D links based on a metric and then find a threshold to divide all the links into two subsets such that D2D links in one subset are set to work on the D2D mode while those in another subset will work on the cellular mode.

Intuitively, a D2D link i should work on a mode that can lead to relatively high channel gains, which hopefully can result in low power consumption. So we use the following channel gain ratio $g(i)$ as the metric to assist mode selection:

$$g(i) = \frac{\frac{\sum_{k=1}^{K} G_{T(i),R(i),k}}{K}}{\frac{\sum_{k \in \{1,\dots,K\} \setminus Q^{\text{legacy}}} G_{T(i),BS,k}}{K-|Q^{\text{legacy}}|}}. \tag{3.1}$$

Basically, $g(i)$ is the ratio of the average channel gain in the D2D mode to that in the cellular mode. Note that $g(i)$ is the ratio between two channel gains, which is different from channel gain. The higher this ratio is, the more likely the link should work on the D2D mode. The hard part is to determine a threshold for this metric to split the D2D links into two modes. Our algorithm performs a linear search on the channel gain ratios of all D2D links and selects the one that leads to minimal total power consumption as the threshold (lines 4–10 in Algorithm 15). We formally present this algorithm as Algorithm 15.

This algorithm uses a subroutine to determine the channel-power assignment $\mathbf{p}$ based on given mode selection $\mathbf{m}$ (line 7 in Algorithm 15). The channel-power assignment subproblem is to determine the sub-channels allocated to each D2D link and the corresponding power assignment. The goal is to minimize total power consumption based on the given mode selection. P is the total power consumption for channel-power assignment $\mathbf{p}$.

The channel-power assignment subroutine is formally presented as Algorithm 16.

Since equations (2.7) and (2.8) are step functions, the channel-power allocation problem still cannot be solved optimally after we are given the mode selection $\mathbf{m}$ for all D2D links. We propose a waterfilling-like algorithm (Algorithm 16), which increases only one D2D link's

Algorithm 15: Joint Algorithm 1 (Joint-1)

Input : $\mathbf{\Gamma} =< \Gamma_i >$, $\mathbf{G} =< G_{T(i),R(j),k} >$
Output: $\mathbf{m} =< m_i >$, $\mathbf{p} =< p_{i,k} >$, $P_{\min}$
Sort all D2D links in the ascending order of channel gain ratio g_i (Eq. 3.1)
and store their indices in an array A;
$j := 0$;
while $j \leq N$ **do**
    $m_{A[i]} := 0$, $i \leq j$ and $i \in [1, ..., N]$;
    $m_{A[i]} := 1$, $j < i < N$ and $i \in [1, ..., N]$;
    $< \mathbf{p}, P >:=$ Set-Channel-Power$(\mathbf{m}, \mathbf{G}, \mathbf{\Gamma})$;
    if $j = 0$ ***or*** $P < P_{\min}$ **then**
        $< \mathbf{m}_{\text{opt}}, \mathbf{p}_{\min}, P_{\min} >:=< \mathbf{m}, \mathbf{p}, P >$;
    $j := j + 1$;
return $< \mathbf{m}_{opt}, \mathbf{p}_{\min}, P_{\min} >$;

data rate by one level at each step, while minimizing total incremental power consumption (lines 3–18 in Algorithm 16).

We find that if mode selection $\mathbf{m}$ and channel rate assignment $\mathbf{r}$ are given, then the channel-power assignment $\mathbf{p}$ can be obtained by solving a Linear Programming (LP) problem, which can be done in polynomial time. We use $P_{\mathbf{m},\mathbf{r}}$ to denote the total power consumption, and use $\mathbf{p}_{\mathbf{m},\mathbf{r}}$ to denote the channel-power allocation solution when channel rate assignment is $\mathbf{r}$ and mode selection solution is $\mathbf{m}$. We formally present the LP for channel-power assignment in the following:

LP-Channel-Power $(\mathbf{m}, \mathbf{r})$

$$P = \min_{<\mathbf{p}>} \sum_{i=1}^{N} \sum_{k=1}^{K} p_{i,k} \tag{3.2}$$

Subject to:

$$\frac{p_{i,k} G_{T(i),R(i),k}}{\sum_{j \neq i} p_{j,k} G_{T(j),R(i),k} + I^{\text{legacy}}_{R(i),k} + N_0} \geq C^{-1}(r_{i,k}), m_i = 1, \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\}; \tag{3.3}$$

$$\frac{p_{i,k} G_{T(i),BS,k}}{\sum_{j \neq i} p_{j,k} G_{T(j),BS,k} + N_0} \geq C^{-1}(r_{i,k}), m_i = 0, \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\} \backslash Q^{\text{legacy}}; \tag{3.4}$$

Algorithm 16: Set-Channel-Power

Input : $\mathbf{m} =< m_i >$, $\mathbf{G} =< G_{T(i),R(j),k} >$, $\mathbf{\Gamma} =< \Gamma_i >$

Output: $\mathbf{p} =< p_{i,k} >$, $P = \sum_{i=1}^{N} \sum_{k=1}^{K} p_{i,k}$

$r_{i,k} := 0, \forall i \in [1, ..., N], \forall k \in [1, ..., K]$;

while 1 **do**

    for *each pair* (i, k) *with* $p_{i,k}$ *not setting to* 0 **do**

        if *Eq. (2.2) is not satisfied* **then**

            Increase $r_{i,k}$ one rate level up;

            $r'_{i,k} := r_{i,k} + \Delta r_{i,k}$;

            $P_{\mathbf{m},\mathbf{r}'} :=$ LP-Channel-Power$(\mathbf{m}, \mathbf{r}')$;

            $P_{\mathbf{m},\mathbf{r}} :=$ LP-Channel-Power$(\mathbf{m}, \mathbf{r})$;

            if *LP-Channel-Power*$(\mathbf{m}, \mathbf{r}')$ *infeasible* **then**

                $W_{i,k} := -1$;

            else

                Calculate $W_{i,k}$ using Eq. (3.8);

        else

            $W_{i,k} := 0$;

    $W_{\max} := \max_{i \in [1,...,N], k \in [1,...,K]} W_{i,k}$;

    $< \mathbf{r}_{\max}, \mathbf{p}_{\max} >:= \underset{<\mathbf{r},\mathbf{p}>}{\operatorname{argmax}} W_{i,k}$;

    if $W_{\max} > 0$ **then**

        $< \mathbf{r}, \mathbf{p} >:=< \mathbf{r}_{\max}, \mathbf{p}_{\max} >$;

        Set some $p_{i,k} := 0$ according to Constraints (2.4);

    else if $W_{\max} = 0$ **then**

        break;

    else

        return $< null, -1 >$;

$P := \sum_{i=1}^{N} \sum_{k=1}^{K} p_{i,k}$;

return $< \mathbf{p}, P >$;

$$\sum_{i=1}^{N} m_i p_{i,k} G_{T(i),BS,k} \leq P^{\text{legacy}}, \forall k \in Q^{\text{legacy}}; \tag{3.5}$$

$$(1 - m_i) p_{i,k} = 0, \forall i \in \{1, \cdots, N\}, \forall k \in Q^{\text{legacy}}; \tag{3.6}$$

$$\sum_{k=1}^{K} p_{i,k} \leq P^{\max}, \forall i \in \{1, \cdots, N\}; \tag{3.7}$$

where $C^{-1}(r_{i,k})$ gives the SINR value corresponding to $r_{i,k}$ for link i and sub-channel k. This LP problem can be efficiently solved in polynomial time. In the simulation, we used the Gurobi Optimizer [36] to solve all LP problem instances.

Next, we explain the structure of Algorithm 16. Initially, the algorithm sets data rates of all link-channel pairs to 0 (line 1). In the while loop, the algorithm tries to find the most power-efficient upgrade in each iteration, which increases the data rate of a link-channel pair one level up (lines 3–18). In the for loop (lines 3–14), the algorithm examines all possible link-channel pairs to find the best one by solving a series of LP-Channel-Power. We use the following *rate-power ratio* to measure power efficiency:

$$W_{i,k} = \frac{\Delta r_{i,k}}{\Delta P_{\mathbf{m},\mathbf{r}}(\Delta r_{i,k})}, \tag{3.8}$$

where $\Delta r_{i,k}$ is the incremental data rate and $\Delta P_{\mathbf{m},\mathbf{r}}(\Delta r_{i,k})$ gives the corresponding incremental power consumption. The algorithm keeps selecting the most power-efficient link-channel pair (according to the rate-power ratio) to upgrade its rate in each iteration (lines 15–18) till the corresponding data rate requirement on each D2D link is satisfied (lines 20–21). To avoid violating constraints (2.4) after a link-channel rate ($r_{i,k}$) is upgraded (line 19), the algorithm disregards some link-channel pairs by setting their power assignments $p_{i,k} := 0$. Setting $p_{i,k} := 0$ ensures that the subchannel k cannot be used by D2D link i working in the cellular mode since some other link j working in the cellular mode has already used it.

The time complexity of Joint-1 (Algorithm 15) is dominated by the while loop, which takes $O(N \cdot T_2)$ time, where T_2 is the running time of Algorithm 16. Similarly, the running time of Algorithm 16 is also dominated by a while loop, which takes $O(N^2K^2L \cdot T_{\text{LP-Channel-Power}})$ time, where L is the number of SNR levels and $T_{\text{LP-Channel-Power}}$ is the time for solving the LP. So the overall time complexity of Joint-1 is $O(N^3K^2L \cdot T_{\text{LP-Channel-Power}})$. Note that in practice, the LP can be solved very quickly using a well-designed LP solver such as the Gurobi Optimizer [36].

As introduced in [50], in reality, BS/UEs broadcast reference signals in a cell every 8ms. After receiving the signal, each UE will report the Channel Quality Identifier (CQI) indicating its channel condition such that the BS can obtain necessary input information for the proposed algorithms. The BS is expected to run a resource optimization algorithm every several seconds, which is certainly doable for the proposed algorithms.

3.3.2 Joint Algorithm 2 (Joint-2)

In this algorithm (denoted as *Joint-2*), we also jointly determine mode selection, channel allocation and power assignment but in a way different from above. Specifically, we enumerate all link-mode-channel triplets; then for each link-mode-channel triplet, we try to find the best power assignment, which, however, is hard to determine without knowing the data rate the corresponding triplet should work at. There are multiple levels for the data rate. Unlike in Joint-1, we employ a greedy approach by pushing the data rate of a link-mode-channel triplet to the highest possible level. Doing so leads to using relatively small number of subchannels over each link, which hopefully causes limited interference to other links, thereby resulting in less power for compensating interference. Joint-2 is formally presented as Algorithm 17.

As mentioned above, if mode selection and channel rate assignment are given, then the corresponding power assignment can be obtained by solving an LP problem, which can be done in polynomial time. Here, similar to LP-Channel-Power, we try to determine the power assignment by solving an LP problem. However, different from Joint-1, in which modes of all links are determined before solving LP-Channel-Power, Joint-2 tries to determine the power assignment for a triplet (instead of a link-channel pair) when some triplets have not yet been

---

Algorithm 17: Joint Algorithm 2 (Joint-2)

Input : $\mathbf{\Gamma} =< \Gamma_i >$, $\mathbf{G} =< G_{T(i),R(j),k} >$
Output: $\mathbf{m} =< m_i >$, $\mathbf{p} =< p_{i,m,k} >$, $P_{\min}$

$r_{i,m,k} := 0, \forall m \in [0,1], \forall i \in [1,...,N], \forall k \in [1,...,K]$;
while 1 **do**
    $W_{i,m,k} := -1, \forall m \in [0,1], \forall i \in [1,...,N], \forall k \in [1,...,K]$;
    for *each triplet* (i,m,k) *with* $p_{i,m,k}$ *not set to* 0 **do**
        $\mathbf{m}' := \mathbf{m}$;
        if *Eq. (2.2) is not satisfied* **then**
            Push $r_{i,m,k}$ to r_{highest} except that
            lower rate level can satisfy Γ_i;
            $m'_i := m$; $r'_{i,m,k} := r_{i,m,k} + \Delta r_{i,m,k}$;
            while $r'_{i,m,k} \geq r_{lowest}$ **do**
                $P_{\mathbf{m}',\mathbf{r}'} :=$ LP-Mode-Channel-Power$(\mathbf{m}', \mathbf{r}')$;
                $P_{\mathbf{m}',\mathbf{r}} :=$ LP-Mode-Channel-Power$(\mathbf{m}', \mathbf{r})$;
                if *LP-Mode-Channel-Power*$(\mathbf{m}', \mathbf{r}')$ *infeasible* **then**
                    Decrease $r'_{i,m,k}$ one rate level down;
                else
                    Calculate $W_{i,m,k}$ using Eq. (3.15);
                    break;
        else
            $W_{i,m,k} := 0$;
    $W_{\max} := \max_{m \in [0,1], i \in [1,...,N], k \in [1,...,K]} W_{i,m,k}$;
    $< \mathbf{m}_{\max}, \mathbf{r}_{\max}, \mathbf{p}_{\max} > := \underset{<\mathbf{m},\mathbf{r},\mathbf{p}>}{\operatorname{argmax}} W_{i,m,k}$;
    if $W_{\max} > 0$ **then**
        $< \mathbf{m}, \mathbf{r}, \mathbf{p} > := < \mathbf{m}_{\max}, \mathbf{r}_{\max}, \mathbf{p}_{\max} >$;
        Set some $p_{i,m,k} := 0$ according to Constraints (3.16)(3.17);
    else if $W_{\max} = 0$ **then**
        break;
    else
        return $< null, null, -1 >$;
$P := \sum_{i=1}^{N} \sum_{m=0}^{1} \sum_{k=1}^{K} p_{i,m,k}$;
return $< \mathbf{m}, \mathbf{p}, P >$;

---

considered (i.e., modes/channels/rates of some links have not yet been determined). We formally present the LP for the power assignment in the following:

LP-Mode-Channel-Power $(\mathbf{m}, \mathbf{r})$

$$P = \min_{<\mathbf{p}>} \sum_{i=1}^{N} \sum_{m=0}^{1} \sum_{k=1}^{K} p_{i,m,k} \tag{3.9}$$

Subject to:

$$\frac{p_{i,1,k} G_{T(i),R(i),k}}{\sum_{j \neq i} p_{j,1,k} G_{T(j),R(i),k} + I^{\text{legacy}}_{R(i),k} + N_0} \geq C^{-1}(r_{i,1,k}), \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\}; \tag{3.10}$$

$$\frac{p_{i,0,k} G_{T(i),BS,k}}{\sum_{j \neq i} p_{j,0,k} G_{T(j),BS,k} + N_0} \geq C^{-1}(r_{i,0,k}), \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\} \backslash Q^{\text{legacy}}; \tag{3.11}$$

$$\sum_{i=1}^{N} m_i p_{i,1,k} G_{T(i),BS,k} \leq P^{\text{legacy}}, \forall k \in Q^{\text{legacy}}; \tag{3.12}$$

$$(1 - m_i) p_{i,0,k} = 0, \forall i \in \{1, \cdots, N\}, \forall k \in Q^{\text{legacy}}; \tag{3.13}$$

$$\sum_{m=0}^{1} \sum_{k=1}^{K} p_{i,m,k} \leq P^{\max}, \forall i \in \{1, \cdots, N\}; \tag{3.14}$$

Similarly, in the above formulation, $\mathbf{p_{m,r}}$ denotes the power assignment corresponding to mode selection $\mathbf{m}$ and channel rate assignment $\mathbf{r}$. Equations (3.9)-(3.14) are similar to equations (3.2)-(3.7) respectively. The difference is that variables $< p_{i,k} >$ are replaced by $< p_{i,m,k} >$ since we consider link-mode-channel triplet in Joint-2 instead of link-channel pair in Joint-1. The objective is to minimize total power consumption. Note that this LP is solved iteratively. Every time when it is solved, for those triplets that have not yet be considered, the correspond rates $(r_{i,m,k})$ are set to 0, resulting in $C^{-1}(r_{i,m,k}) = 0$; while for those triplets that have been considered, their rates are set to the values determined in previous steps.

Next, we explain the structure of Algorithm 17. Initially, the algorithm sets the data rates of all link-mode-channel triplets to 0 (line 1). In the outer while loop, the algorithm tries to find the most power-efficient upgrade in each iteration, which increases the data rate of a link-mode-channel triplet to the highest possible level (lines 4–23). In the for loop (lines 3–19), the algorithm examines all possible link-mode-channel triplets to find the best one by solving a series of LP-Mode-Channel-Power. Similarly, we use the following *rate-power ratio* to measure power efficiency:

$$W_{i,m,k} = \frac{\Delta r_{i,m,k}}{\Delta P_{\mathbf{m},\mathbf{r}}(\Delta r_{i,m,k})}, \tag{3.15}$$

where $\Delta r_{i,m,k}$ is the incremental data rate and $\Delta P_{\mathbf{m},\mathbf{r}(\Delta r_{i,m,k})}$ gives the corresponding incremental power consumption. Joint-2 keeps selecting the most power-efficient triplet (i, m, k) (according to equations (3.15)), and pushes its rate to the highest possible level by checking if the above LP-Mode-Channel-Power can still return a feasible solution (lines 7–17). By doing so, the algorithm contributes to the rate requirement of the corresponding link in a greedy manner. This procedure stops when the rate requirement of every link is satisfied (lines 25–26).

After each rate upgrade for some link, the algorithm disregards those link-mode-channel triplets with conflicts to ensure feasibility by setting the corresponding $p_{i,m,k} := 0$. Specifically, the power assignment for some link-mode-channels ($p_{i,m,k}$) need to be set to 0 in order to avoid violating constraints (3.16) and (3.17) as listed below (line 24):

$$p_{i,0,k}p_{i,1,k} = 0, \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\}; \tag{3.16}$$

$$(1 - m_i)p_{i,0,k} \sum_{j \neq i} (1 - m_j)p_{j,0,k} = 0, \forall i, j \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, K\} \backslash Q^{\text{legacy}}. \tag{3.17}$$

Constraints (3.16) ensure that each D2D link can work in only one mode, either cellular mode or D2D mode. Constraints (3.17) make sure that any two D2D links both working in cellular mode cannot share a common sub-channel.

The time complexity of Joint-2 (Algorithm 17) is dominated by the outer while loop. The running time of each iteration is dominated by the for loop, which takes $O(NKL \cdot$

$T_{\text{LP-Mode-Channel-Power}}$) time, where L is the number of SNR levels and $T_{\text{LP-Mode-Channel-Power}}$ is the time for solving the LP. Thus, the overall time complexity of Joint-2 is $O(N^2K^2L \cdot T_{\text{LP-Mode-Channel-Power}}))$. Again, the LP can be solved very quickly in practice.

3.4 Performance Evaluation

In this section, we present and analyze simulation results to justify effectiveness of the proposed algorithms.

In the simulation, the coverage region of the cell was a disk with a radius of $R = 300$m. A BS was located at the center of the cell, and N^{legacy} legacy users were randomly distributed in the cell. $Q^{\text{legacy}} = 2 \times N^{\text{legacy}}$ sub-channels have been randomly assigned to legacy users. For each pair of D2D link $T(i), R(i)$, the receiver $R(i)$ was randomly placed in the circle centered at the sender $T(i)$ with a radius of $D_{\max}$, which follows a 2D uniform distribution. For each D2D link i, the data rate requirement Γ_i was randomly chosen, which follows a uniform distribution between $\Gamma_{\min}$ and $\Gamma_{\max}$. In order to guarantee the QoS of legacy users, the aggregated interference on each legacy sub-channel from D2D links cannot exceed a threshold P_{legacy}. If a D2D link works in the cellular mode where data traffic is relayed by the BS, then it cannot use the legacy sub-channels reserved for legacy users. The sub-channel gains were set to follow the free space model [29]:

$$G = (20 \log_{10}(d) + 20 \log 10(f) + 92.45)(1 + \sigma), \tag{4.1}$$

where d is the distance between transmitter and receiver in the unit of km and f is the center frequency in the unit of GHz. σ is a zero mean random variable following standard distribution. We summarize common simulation settings in the Table 3.2.

As mentioned above, the link data rate is an increasing step function of its SNR levels. According to the IEEE 802.16e standard [43], we show how we set per-channel link data rates using Table 3.3. All the values presented here are calculated based on the settings that

Table 3.2: Common Simulation Settings

Parameter	**Value**
Radius of the cell	300m
Sub-channel bandwidth	0.4MHz
Background noise	-85dBm
Max transmit power ($P^{\max}$)	25mW
Gauss variance of σ	0.5
Min data rate requirement ($\Gamma_{\min}$)	0.4Mbps
P^{legacy}	-87.21dBm
No. of sub-channels for each legacy user	2
Frequency band (f)	1.92GHz

the sub-channel bandwidth is 0.4MHz and the antenna gain is 2dBi. Note that link data rate is a linear function of the sub-channel bandwidth, therefore we can easily obtain a similar step function if we are given a different sub-channel bandwidth.

Modulation	**Code Rate**	**Min SNR (dB)**	**Rates(Mbps)**
QPSK	1/2	10	0.4
16QAM	1/2	14.5	0.8
16QAM	3/4	17.25	1.2
64QAM	2/3	21.75	1.6
64QAM	3/4	23	1.8

Table 3.3: SNR thresholds and the corresponding per-channel data rates according to the WiMAX standard [43]

In the simulation, we compared the proposed algorithm with the following baseline algorithms:

1. All D2D links in the cellular mode with random sub-channel allocation (*All-Cellular*): In this algorithm, all D2D links work in the cellular mode and sub-channels are randomly allocated to D2D links such that each D2D link gets the same number of sub-channels.

2. All D2D links in the D2D mode with random sub-channel allocation (*All-D2D*): In this algorithm, all D2D links work in the D2D mode and sub-channels are randomly allocated to D2D links such that each D2D link gets the same number of sub-channels.

3. Random mode selection and random sub-channel allocation algorithm (*Random*): Each D2D link's mode is randomly determined, with 50% probability for each mode. Channel allocation is the same as that of the other baseline algorithms.

Note that in all these three baseline algorithms, after random channel allocation, they assign power to each sub-channel using a greedy subroutine: start channel-power assignment from certain level such that the link can have the highest possible SNR (that can lead to the highest data rate); lower channel-power assignment as long as the corresponding link data rate is large enough to meet the given requirement.

We compared the proposed joint algorithms against the three baseline algorithms in terms of total power consumption using the following 5 scenarios:

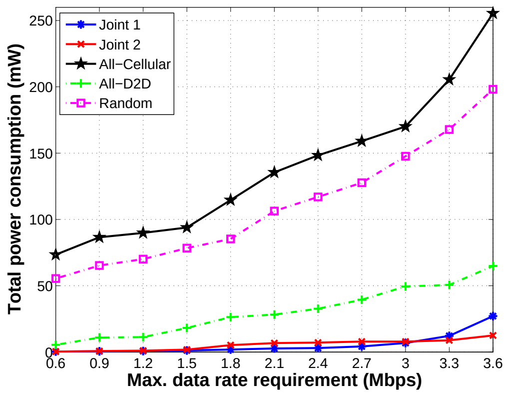
1. Scenario 1: We changed the maximum rate requirement $\Gamma_{\max}$ from 0.6Mbps to 3.6Mbps with a step size of 0.3Mbps. The other parameters were set as follows: $N = 12$, $D_{\max} = 15$m, $K = 34$ and $N^{\text{legacy}} = 5$.

2. Scenario 2: We increased the number of D2D links N from 4 to 24 with a step size of 2. The other parameters were set as follows: $\Gamma_{\max} = 1.8$Mbps, $D_{\max} = 15$m, $K = 34$ and $N^{\text{legacy}} = 5$.

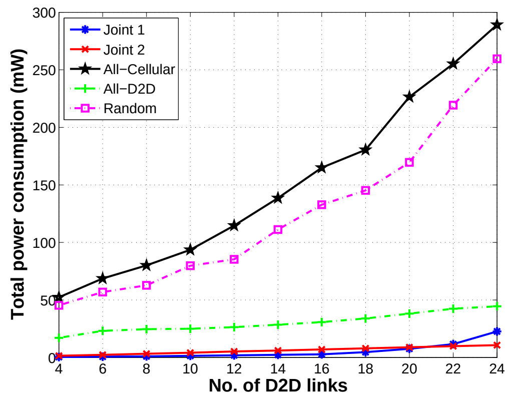
3. Scenario 3: We varied the maximum distance of D2D links $D_{\max}$ from 5m to 40m with a step size of 5m. The other parameters were set as follows: $\Gamma_{\max} = 1.8$Mbps, $N = 12$, $K = 34$ and $N^{\text{legacy}} = 5$.

4. Scenario 4: We increased the number of available sub-channels K from 22 to 50 with a step size of 4. The other parameters were set as follows: $\Gamma_{\max} = 1.8$Mbps, $N = 12$, $D_{\max} = 15$m and $N^{\text{legacy}} = 5$.

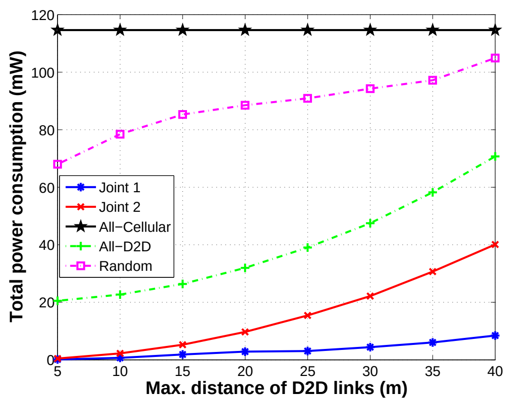
5. Scenario 5: We increased the number of legacy users N^{legacy} from 2 to 10 with a step size of 1. The other parameters were set as follows: $\Gamma_{\max} = 1.8$Mbps, $N = 12$, $D_{\max} = 15$m and $K = 34$.



(a) Max. data rate requirement ($\Gamma_{\max}$)



(b) No. of D2D links (N)

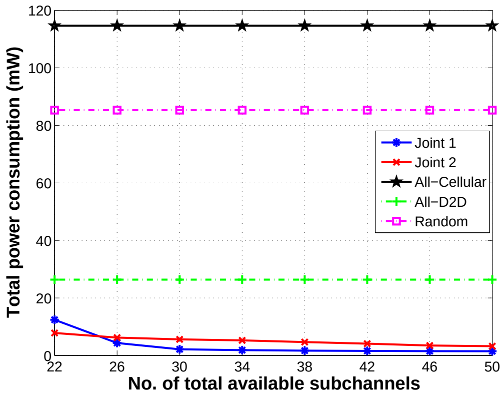


(c) Max. distance of D2D links ($D_{\max}$)

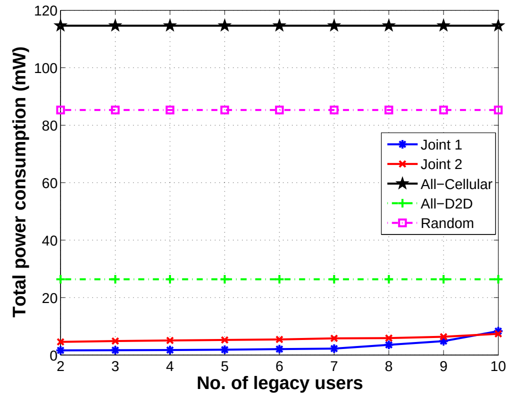
Figure 3-2: Impact of $\Gamma_{\max}$, N and $D_{\max}$ on the total power consumption

The simulation results are presented in Figs. 3-2 and 3-3. We can make the following observations from these results:

1) In all scenarios, the proposed joint algorithms consistently outperform the baseline algorithms. On average, Joint-1 and Joint-2 achieve 88% and 84% power savings compared to All-Cellular, respectively. This shows that D2D communications can significantly reduce power consumption compared to the traditional communication approach. Moreover, compared to All-D2D, the proposed algorithms can lead to an average of 78% and 68% power savings, respectively. Compared to Random, the proposed algorithms result in an average of 86% and 82% power savings, respectively. This justify our claim that when using D2D com-



(a) No. of total available sub-channels (K)



(b) No. of legacy users (N^{legacy})

Figure 3-3: Impact of K and N^{legacy} on the total power consumption

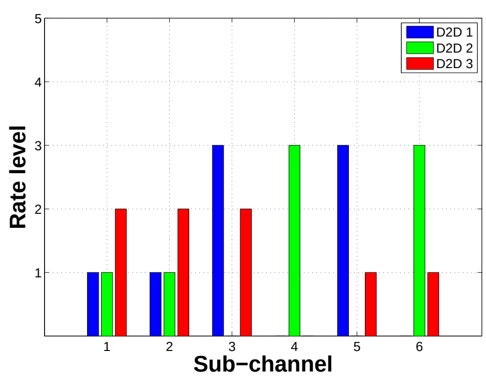
munications, mode selection, channel allocation and power assignment need to be carefully determined.

2) From Figs. 3-2(a) and 3-2(b), we can see that no matter which algorithm is used, the total power consumption increases monotonically with the data rate requirement and the number of D2D links. However, the proposed joint algorithms are superior to the baseline algorithms since unlike them, the corresponding power consumption grows very slowly with these two important parameters. This shows that compared to simple 3-step greedy methods, joint decision making along with LP-based optimization can lead to significant performance improvement.

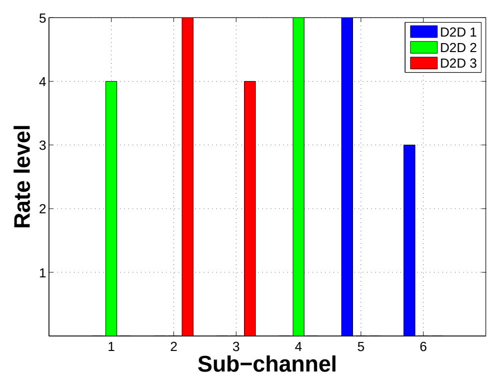
3) From Fig. 3-2(c), we can see that a longer D2D link distance leads to more power consumption for both our joint algorithms and baseline algorithms except All-Cellular where all D2D links work in the cellular mode so that they have nothing to do with this parameter. Since all D2D links have to maintain their received SNR at certain levels in order to meet their data rate requirements, longer D2D link distances will result in higher transmit power for those D2D links working in the D2D mode. Power consumption of Joint-2 grows faster than that of Joint-1. This is because usually rate levels on active sub-channels obtained from Joint-2 are higher than those from Joint-1, thus, higher transmit power is necessary to

achieve higher SNRs in Joint-2.

4) From Figs. 3-3(a) and 3-3(b), we can make two interesting findings. First, power consumption given by all the baseline algorithms remains the same even with more sub-channels. Power assignment in the baseline algorithms uses a simple greedy procedure. If a link's data rate requirement can be satisfied by certain number of sub-channels then the algorithms will not use more sub-channels. In other words, more sub-channels do not necessarily lead to better performance for those baseline algorithms. Second, the proposed joint algorithms result in less power consumption with more sub-channels. That is because our algorithms always select the most power-efficient sub-channel to use in each step. More available sub-channels means the algorithms has more options to choose from. If there are better sub-channels from the additional set of available sub-channels, total power consumption given by our algorithms will be reduced. Otherwise they remain the same just like the baseline algorithms. This again shows that joint decision making with LP-based optimization outperforms simple greedy methods.



(a) Rate level on each sub-channel by Joint-1



(b) Rate level on each sub-channel by Joint-2

Figure 3-4: Illustration of channel-rate allocation by Joint-1 and Joint-2

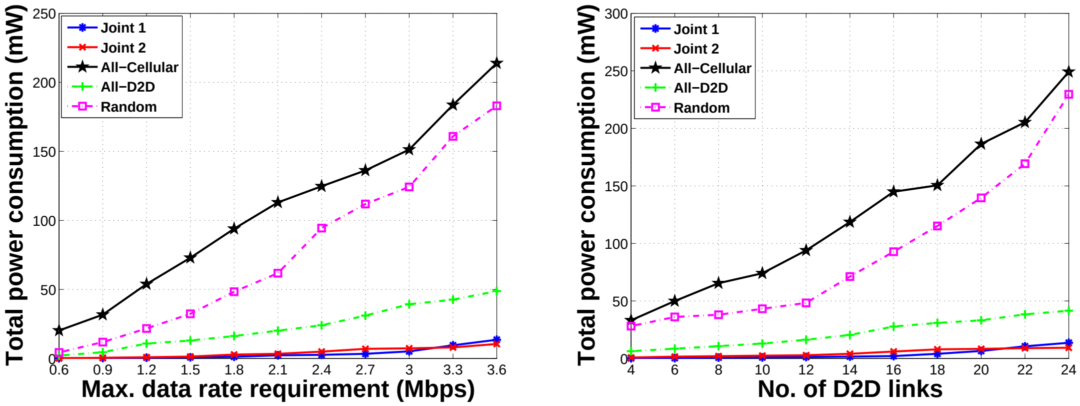
5) In some cases, power consumption of Joint-1 is less than that that of Joint-2. However, in some cases (e.g. $\Gamma_{\max} \geq 3.0$Mbps, $N \geq 20$, $K \leq 26$, and $N^{\text{legacy}} \geq 9$ in Figs. 3-2 and 3-3), Joint-2 offers nearly the same or better performance. Usually, Joint-2 uses less sub-

channels but higher rate levels, while Joint-2 uses more sub-channels but lower rate levels. Thus, in those cases with limited available sub-channels, Joint-2 may likely perform better than Joint-1. To further demonstrate this, we conducted simulation with a small case with $\Gamma_{\max} = 3.5$Mbps, $N = 3$ and $K = 6$ and show the corresponding results in Fig. 3-4. We can clearly see that differences between Joint-1 and Joint-2 on sub-channel utilization.

Modulation	**Code Rate**	**Min SNR (dB)**	**Rates(Mbps)**
QPSK	1/12	-6.50	0.06
QPSK	1/9	-4.00	0.092
QPSK	1/6	-2.60	0.15
QPSK	1/3	-1.00	0.24
QPSK	1/2	1.00	0.352
QPSK	3/5	3.00	0.472
16QAM	1/3	6.60	0.592
16QAM	1/2	10.00	0.764
16QAM	3/5	11.40	0.964
64QAM	1/3	11.80	1.092
64QAM	1/2	13.00	1.328
64QAM	3/5	13.80	1.560
64QAM	3/4	15.60	1.808
64QAM	5/6	16.80	2.048
64QAM	11/12	17.60	2.220

Table 3.4: SNR thresholds and the corresponding per-channel data rates according to the LTE standard [50]

In addition, we also evaluated the performance of the proposed algorithms based on the LTE standard. As in [50], we used the following settings in our simulation, which are listed in Table 3.4. The other settings were the same as those in scenarios 1 and 2. The corresponding results are shown in Fig. 3-5, from which we can make the following observations: Similarly, we can see that the proposed joint algorithms consistently outperform the baselines. On average, Joint-1 and Joint-2 achieve 98% and 97% power savings compared to All-Cellular respectively. This shows that D2D communications can significantly reduce transmit power consumption on UEs compared to the traditional approach in LTE networks. Moreover, we can make similar observations when we compare Joint-1 and Joint-2 with All-D2D and Random. Specifically, compared to All-D2D, the proposed algorithms lead to an average



(a) Max. data rate requirement (Γ_{max})   (b) No. of D2D links (N)

Figure 3-5: Impact of $\Gamma_{\max}$ and N on the total power consumption according to the LTE standard [50]

of 88% and 84% power savings, respectively; and compared to Random, they result in an average of 96% and 95% power savings, respectively.

Chapter 4

Base Station Consolidation in Virtualized Cognitive Radio Networks

4.1 Overview

Virtualization is the creation of a virtual (rather than actual) version of certain physical resources, such as a computer, storage device, or network resources. Virtualization has emerged as a useful technology for improving resource utilization and power efficiency. For example, in a virtualized data center, Virtual Machines (VMs) can be created to host applications and servers can be consolidated by migrating VMs such that idle servers and chassis, can be shut down or put into sleep. Virtualization technology has been introduced to wireless networking recently [48]. In general, network virtualization enables deploying customized services and resource management solutions in isolated slices on a shared physical network. Particularly, with wireless resource virtualization, multiple Mobile Virtual Network Operators (MVNOs) can be supported over a shared physical wireless network and traffic loads in a BS can be easily migrated to more power-efficient BSs in its neighborhood such that idle BSs can be turned off or put into sleep to save power.

Emerging Cognitive Radio (CR) technology and the Dynamic Spectrum Access (DSA) approach [5] enable unlicensed wireless users (a.k.a secondary users) to sense and access the under-utilized spectrum opportunistically even if it is licensed. CRs have been considered as the next generation wireless communication technology that can lead to better spectrum utilization and higher network capacity.

In this work, we propose to leverage load migration and BS consolidation for green communications and consider a power-efficient network planning problem in virtualized Cognitive Radio Networks (CRNs) with the objective of minimizing total power consumption while meeting traffic load demand of each MVNO. We find that the problem can divided into two subproblems: the channel assignment problem and the load allocation problem. The channel assignment problem seeks a solution that assigns a channel for each BS in an interference-free manner. The load allocation problem is to determine which subset of BSs to turn off (or put into sleep) and how to allocate load of each MVNO on every BS to active BSs. Power savings can be achieved by migrating loads of MVNOs to more power-efficient BSs and/or shutting down BSs to save idle power. Note that since different CR BSs work on different channels, their power efficiency might be different because to maintain certain transmission range, the BS using a low-frequency channel can use less power than that using a high-frequency channel due to the signal propagation property.

Even though green wireless networking has attracted extensive attention recently, most previous works were focused on 3G/4G/WiFi networks (rather than CRNs) without addressing the case where there are multiple MVNOs in the network. To the best of our knowledge, we are the first to propose to leverage load migration and BS consolidation for green communications in a virtualized CRN (with multiple MVNOs), and present theoretically well-founded and practically efficient algorithms to solve the corresponding optimization problems.

The differences between our work and these related works are summarized as follows: 1) Unlike most papers on wireless resource virtualization which were mainly focused on how to design and implement resource virtualization at one node, we aim to leverage load migration and BS consolidation (that can be enabled by virtualization) for reducing power consumption of the whole network 2) Most previous works on green wireless networking studied power-efficient resource management problems in the context of 3G/4G/WiFi networks rather than CRNs and did not address the case where there are multiple MVNOs in the network. 3) Different from most works on spectrum sharing in CRNs which aimed at improving spectrum utilization and network capacity, we study a power-efficient network planning problem in the context of a network with virtualized CR BSs, which has never been done before.

The rest of this chapter is organized as follows. We present the problem definition in Section 4.2. We propose an optimization framework to guide the algorithm design in Section 4.3. An joint algorithm is proposed and discussed in Section 4.4. At last, we present simulation results in Section 4.5.

4.2 Problem Definition

We summarize major notations in the following table for quick reference.

In this work, we consider a CRN with N BSs, which is shared by K MVNOs. The available spectrum is divided into a set of orthogonal *channels*. Note that since we study a network planning problem rather than a MAC layer channel selection problem, the channel considered here represents a relatively large portion of the spectrum and may include a group of sub-channels defined in the context of OFDMA. There are a set H_i of channels available to each BS i, which may change over time. Channel availability information can be obtained from a spectrum database (as suggested by FCC) or using a spectrum sensing method [87]. Since available channels of a CR may be distributed over a large range of spectrum, they may have (or be able to support) quite different properties such as channel gain, data rate, etc. Each BS is assigned one channel to support wireless users of multiple MVNOs for a certain period of time.

With virtualization, wireless resources (such as timeslots in an OFDMA frame) are allo-

Table 4.1: Major Notations

Variable	Description
A_i	Idle power
B_i^f	The load coefficient in the power consumption function of BS i
C_i^h	The capacity of BS i on channel h
H_i	A set channels available to BS i
$H_{\max}$	The maximum number of available channels on a BS
K	The number of MVNOs
L_{ik}	The load demand of MVNO k on BS i
l_i	The total traffic load on BS i
N	The number of BSs
S_i	The neighbor set of BS i
I_i	The interference set of BS i
x_i	Decision variable: $x_i = 1$ if BS i is turned on; 0, otherwise.
y_i^h	Decision variable: $y_i^h = 1$ if channel h is assigned to BS i; 0, otherwise.
l_{ik}^{jh}	Decision variable: The amount of load of MVNO k on BS i that is migrated to BS j on channel h.

cated dynamically to slices, each of which may correspond an MVNO or a group of service flows of an MVNO. In a BS, a slice manager [48] (similar to hypervisor in the context of server virtualization) can be used to manage resource allocation for slices with the goal of achieving isolation, customization and efficient utilization of resources. Each MVNO provides wireless communication services for a group of mobile users, which create certain traffic loads on each BS. We use a $N \times K$ matrix L to specify the traffic load of MVNO k on BS i. Each BS i can migrate part of or all of its traffic load to a set S_i of neighboring BSs. With virtualization, a BS can quickly adjust resource allocation for its slices to accommodate traffic loads migrated from other BSs. Different (conservative or aggressive) criteria can be used to identify such a neighbor set for each BS. Note that $i \in S_i$. In addition, for each BS i, there is a set of BSs I_i which can potentially interfere with BS i if they work on the same channel. If channel h is assigned to BS i, then any BS in I_i cannot be assigned channel h.

Similarly, different criteria can be used to identify such an interference set for each BS. Note that $i \notin I_i$ (this is just a technical agreement for easy presentation). In the simulation, we used relatively conservative methods to identify these two sets for each BS, which will be explained in Section 4.5.

Similar as in [69], we adopt a simple and widely-used linear model for power consumption of a BS i, which is given as follows:

$$P_i(l_i, f) = A_i + B_i^f * l_i, \tag{2.1}$$

where A_i is a constant that specifies the idle power usage, l_i is the traffic load, and B_i^f is the load coefficient. Note that the value of this coefficient may vary with transmission frequency f because usually low-frequency wireless signals travel longer than high-frequency signals if transmitted at a given power level, in other words, to maintain certain transmission range, the BS using a low-frequency channel can use less power than that using a high-frequency channel. However, the values may be the same for multiple channels on a common spectrum band. The values of A_i and B_i^f can be obtained via a profile-based approach such as that in [69] or estimated using a signal propagation model [70]. We say a BS is *power-proportional* if $A_i = 0$. Currently, almost no BS is power-proportional. However, we still consider this special case since with the advancement of communication hardware and cooling technology, it might be possible to significantly reduce idle power to make it close to zero in the future.

We are interested in finding a resource allocation solution that specifies which channel to be assigned to each BS and how to allocate traffic loads of each MVNO. A resource allocation solution is said to be *feasible* if available channels are assigned to BSs in a interference-free manner, the traffic load demands specified by the matrix L can be satisfied and the total load on each BS does not exceed its capacity. Now, we are ready to define the optimization problem.

Definition 10. Given N BSs, K MVNOs, a $N \times K$ traffic load matrix L and the set H_i of available channels on each BS i, $i \in \{1, \cdots, N\}$, a power-efficient network planning problem seeks a feasible resource allocation solution that minimizes the total power consumption of BSs.

Note that this network planning procedure can be conducted on a relatively large time scale, e.g., 30 minutes or 1 hour. In addition, we are only interested in finding out how to distribute traffic loads of MVNOs among BSs in a network, however, how to allocate resources to slices to support multiple MVNOs to meet their load demands in a single BS is out of scope of this work, but has been studied in [48, 52, 89].

This optimization problem is very hard to solve since its subproblem, interference-free channel assignment, is known to be NP-hard [75]. Therefore, we first present an MILP formulation to provide optimal solutions.

Decision variables:

- $x_i = \{0, 1\}$: $x_i = 1$ if BS i is turned on; 0, otherwise.
- $y_i^h = \{0, 1\}$: $y_i^h = 1$ if channel h is assigned to BS i; 0, otherwise.
- $l_{ik}^{jh} \geq 0$: The amount of load of MVNO k on BS i that is migrated to BS j on channel h.

MILP-Green

$$\min_{\langle x_i, y_i^h, l_{ik}^{jh} \rangle} \sum_{i=1}^{N} (A_i x_i + \sum_{h \in H_i} B_i^h (\sum_{k=1}^{K} \sum_{j : i \in S_j} l_{jk}^{ih})) \tag{2.2}$$

Subject to:

$$\sum_{h \in H_i} y_i^h \leq x_i, \ \forall i \in \{1, \cdots, N\}; \tag{2.3}$$

$$y_i^h + \frac{\sum_{j \in I_i} y_j^h}{N} \leq 1, \ \forall i \in \{1, \cdots, N\}, \ \forall h \in H_i \tag{2.4}$$

$$\sum_{j \in S_i} \sum_{h \in H_j} l_{ik}^{jh} = L_{ik}, \ \forall i \in \{1, \cdots, N\}, \ \forall k \in \{1, \cdots, K\}; \tag{2.5}$$

$$\sum_{k=1}^{K} \sum_{j: i \in S_j} l_{jk}^{ih} \leq y_i^h C_i^h, \ \forall i \in \{1, \cdots, N\}, \ \forall h \in H_i. \tag{2.6}$$

In this formulation, the objective (2.2) is to minimize total power consumption of BSs. By abusing the notation a little bit, we use h to denote both channel h and its central frequency. As described before, each BS can only be assigned one channel, which is guaranteed by constraints (2.3). Constraints (2.4) ensure that channels are assigned in an interference-free manner, i.e., if channel h is assigned to BS i, then none of BSs in the interference set I_i can be assigned this channel. Each MVNO k has a load demand L_{ik} on each BS i, which must be satisfied by using BS i and/or BSs in the neighbor set S_i via load migrations. This is ensured by constraints (2.5). The last set of constraints (2.6) make sure that each BS i has sufficient capacity to support its own traffic loads and those migrated from neighboring BSs. Note that the capacity of a BS can be conservatively set to certain percentage of its actual capacity to guarantee quality of service since it may need to serve loads migrated from neighboring BSs. It is known that solving such an MILP may take exponentially long time, especially for large cases. Hence, we present polynomial-time algorithms in the next section.

4.3 Optimization Framework

In this section, we first present a 2-step framework to guide algorithm design. Essentially, the optimization problem defined above consists of two subproblems: *channel assignment and load allocation*. The channel assignment subproblem is to determine which channel to be assigned to each BS. The load allocation subproblem is to determine which subset of BSs should be turned on and how to distribute loads to active BSs. We formally present the optimization framework in the following.

Algorithm 18: The Optimization Framework

Step 1 Compute a channel assignment that maximizes the total weight (defined below) of assigned channels;

Step 2 Based on the channel assignment, obtain a minimum-power load allocation.

It is important to obtain a channel assignment that can (hopefully) lead to a minimum power solution to the original problem. Our approach is to set a weight for each channel h on BS i, $w(i,h) = \frac{C_i^h}{P_i(C_i^h,h)}$, where C_i^h is the capacity of BS i on channel h and $P_i(C_i^h, h)$ gives the total power needed to support C_i^h amount of loads on BS i with channel h. Therefore this weight function returns per-watt load that can be supported by assigning channel h to BS i. By finding a channel assignment with maximum total weight, we can (hopefully) have sufficient capacity to accommodate traffic loads with low power consumption. In Step 2, an algorithm can be used to obtain a load allocation based on the channel assignment computed in Step 1.

Next, we present an approximation algorithm for channel assignment. For the load allocation problem, we present a polynomial-time optimal algorithm for a special case where BSs are power-proportional as well as two fast heuristic algorithms for the general case.

4.3.1 Channel Assignment Algorithm

The channel assignment problem is to determine a channel assignment that can result in the maximum total weight (defined above) and ensures that any two BSs that interfere with each other are given two different channels. To assist computation, we construct an auxiliary graph, the Multi-Channel Contention Graph (MCCG), to model conflict (interference) in a network with multiple heterogeneous channels, which was proposed in our previous work [75]. To ensure the completeness of the presentation, we briefly describe how it is constructed. In an MCCG $G_C(V_C, E_C)$, every vertex corresponds to a BS-channel pair in $\mathcal{A}$, where $\mathcal{A} = \{(i,h) : \forall i \in \{1, \cdots, N\}, \forall h \in H_i\}$. There is an undirected edge connecting two vertices in V_C if their corresponding BS-channel pairs *interfere* with each other. Two BS-channel pairs (i,h) and (j,h') are said to interfere with each other if 1) $i = j$ or 2) $h = h'$ and ($j \in I_i$ or $i \in I_j$), where I_i and I_j are the interference sets of BS i and j respectively. Note that there is an undirected edge between every two vertices corresponding to BS-channel pairs that contain a common BS because they always conflict with each other no matter which channel is considered since a BS can only work on one channel. In other words, all vertices corresponding to a common BS form a clique in G_C. This case is covered by the first condition. Now, we are ready to present the channel assignment algorithm.

Algorithm 19: The Channel Assignment Algorithm

Step 1 $V := V_C$; Set all $\langle y_i^h \rangle$:=0;

Step 2 **while** ($V \neq \emptyset$)
  $v_{\max} := \text{argmax}_{v \in V} \frac{w(v)}{\Delta_v + 1}$, where Δ_v is the degree of vertex v on G_C;
  $y_{i_{\max}}^{h_{\max}} := 1$, where $v_{\max} = (i_{\max}, h_{\max})$;
  $V := V - V_{\max}$, where $V_{\max}$ is the set of vertices share a common edge with $v_{\max}$;
endwhile

Step 3 **return** $\langle y_i^h \rangle$;

Every time, the algorithm selects a vertex (BS-channel pair) with the maximum weight-to-degree ratio and assigns the channel accordingly. In this way, vertices with high weights and low interference impact are expected to be selected for channel assignment. We analyze the performance and time complexity of this algorithm in the following.

Theorem 4. Algorithm 19 is a $\frac{1}{\Delta}$-approximation algorithm for the channel assignment problem (where Δ is the maximum vertex degree on the MCCG) and has a time complexity of $O(N^2 H_{\max})$ (where $H_{\max}$ is the maximum number of available channels on a BS).

Proof. Due to the way how the MCCG is constructed, the channel assignment problem can be transformed to the maximum weight independent set problem on an MCCG. It has been shown in [71] that a greedy algorithm that selects a vertex with the maximum weight-to-degree ratio has an approximation ratio of $\frac{1}{\Delta}$ for the maximum weight independent set problem. Hence, our algorithm offers the same approximation ratio for the channel assignment problem.

Every time, it takes $O(NH_{\max})$ to assign channel for one BS. Hence, the overall time complexity of the proposed algorithm is $O(N^2 H_{\max})$. This completes the proof. $\square$

4.3.2 Load Allocation Algorithms

After the channel assignment is determined, we can solve the load allocation subproblem. First, we construct a directed auxiliary graph $G_f(V_O \bigcup V_B \bigcup \{s, z\}, E_f)$ to assist computation. In this graph, each vertex $u \in V_O$ corresponds to a BS-MVNO pair (i, k) and each pair of vertices $v_j^{in}, v_j^{out} \in V_B$ correspond to a BS j. There is a directed edge from each vertex $u = (i, k) \in V_O$ to $v_j^{in} \in V_B$ (corresponding to BS j) if $j \in S_i$, where S_i is the neighbor set of BS i defined above. The cost and capacity of such an edge e are set to $w_e = 0$ and $C_e = \infty$ respectively. Moreover, there is a directed edge from each $v_j^{in} \in V_B$ to $v_j^{out} \in V_B$, whose cost and capacity are set to B_j^h and C_j^h respectively. where B_j^h is the load coefficient in the power consumption function of BS j and C_j^h is the capacity of BS j when using assigned channel h. In addition, we create a virtual sink z and there is a directed edge from each $v_j^{out} \in V_B$ to

z, whose cost and capacity are set to 0 and ∞ respectively. We also create a virtual source s and there is a directed edge from s to each $u \in V_O$, whose cost and capacity are set to 0 and L_{ik} respectively.

Next, we use an example in Fig. 4-1 to show how to construct this graph. In this example, we have 2 MVNOs and 4 BSs which are assigned channels 1, 2, 3 and 4 respectively. In addition, both neighbor sets S_1 and S_2 include BSs 1 and 2.

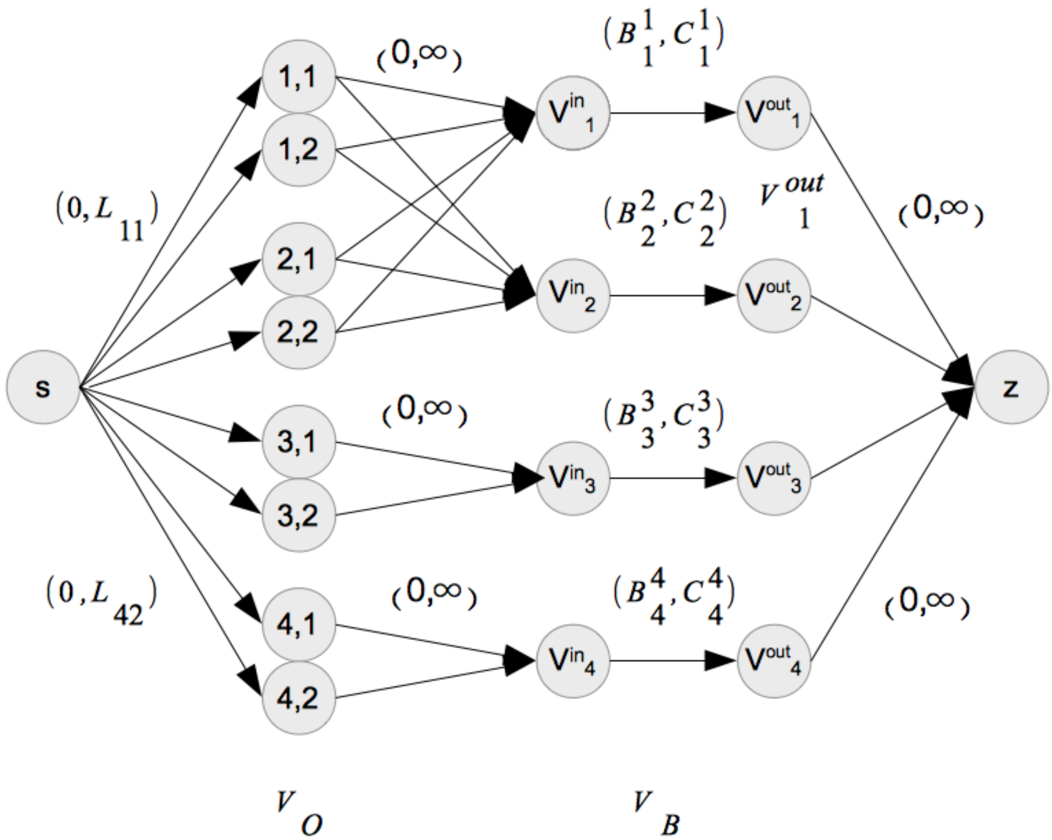


Figure 4-1: An auxiliary flow graph

With this auxiliary flow graph, the load allocation problem can be transformed to a min-cost flow problem in this graph. Specifically, for the power-proportional case where the power consumption of each BS i is $B_i^h * l_i$, then there is no need to turn off any BS and the load allocation problem becomes a traditional min-cost flow problem [2], which can be formulated as the following LP problem, in which f_e specifies the amount of flow over edge e.

LP-Flow-PP:

$$\min_{\langle f_e \rangle} \sum_{e \in E} w_e f_e \tag{3.1}$$

$$\sum_{e \in E_s^{out}} f_e = L_{total}; \tag{3.2}$$

$$\sum_{e \in E_v^{in}} f_e = \sum_{e \in E_v^{out}} f_e, \forall v \in V_O \bigcup V_B; \tag{3.3}$$

$$f_e \leq C_e, \quad \forall e \in E_f. \tag{3.4}$$

In the formulation, w_e and C_e are the cost and capacity of edge e respectively. E_v^{out} and E_v^{out} are the set of edges going out from v and into v respectively. L_{total} is the total load of all MVNOs, i.e., $L_{total} = \sum_{i=1}^{N} \sum_{k=1}^{K} L_{ik}$. We present an algorithm for this special case of load allocation problem in the following.

Algorithm 20: The Min-Cost-Flow-Based Load Allocation Algorithm for the Power-Proportional Case

---

Step 1 Construct the auxiliary flow graph G_f;

Step 2 Find a min-cost $s - t$ flow allocation $\langle f_e \rangle$ on G_f by solving the LP-Flow-PP;

Step 3 **forall** $e = (u, v_j^{in})$, where $u = (i, k) \in V_O$ and $v_j^{in} \in V_B$
$L_{ik}^{jh} = f_e$, where h is the channel assigned
to BS j;
endforall
return $\langle L_{ik}^{jh} \rangle$.

Theorem 5. The min-cost-flow-based algorithm optimally solves the power-proportional case of the load allocation problem in polynomial time.

Proof. Constraint (3.2) ensures the total amount of $s - t$ flow is equal to the total load demand. Since the capacity of each link e going out from s to $u = (i, k) \in V_O$ is set to

$C_e = L_{ik}$, constraint (3.2) ensures the load demand of each MVNO k on every BS i is satisfied. Moreover, due to the way how the capacity of each link e between $v_j^{in}, v_j^{out} \in V_B$ is set (i.e., $C_e = C_j^h$), constraints (3.4) make sure that each BS j has sufficient capacity to support all loads allocated to it. The objective function (3.1) minimizes the total cost of flow, which is equivalent to minimizing the total power consumption since the costs of all edges are set to 0 except that those between $v_j^{in}, v_j^{out} \in V_B$ are set to the load coefficient B_i^h of the power consumption function.

This LP problem has no more than $(N(K+2) + N^2K)$ variables and no more than $(2N(K+2) + N^2K + 1)$ constraints since G_f has $(N(K+2)+2)$ vertices and no more than $(N(K+2) + N^2K)$ edges. Hence, it can be solved in polynomial time. This completes the proof. □

For the general case where there is a non-zero idle power for each BS, the load allocation problem can also be formulated as another flow problem on G_f, in which f_e specifies the amount of flow over edge e; and z_e is an integer decision variable that indicate if edge e is activated ($z_e = 1$) or not ($z_e = 0$).

MILP-Flow:

$$\min_{\langle f_e, z_e \rangle} \sum_{e \in E} (a_e z_e + w_e f_e) \tag{3.5}$$

$$\sum_{e \in E_s^{out}} f_e = L_{total};$$

$$\sum_{e \in E_v^{in}} f_e = \sum_{e \in E_v^{out}} f_e, \ \ \forall v \in V_O \bigcup V_B;$$

$$f_e \leq z_e C_e, \quad \forall e \in E_f^{intra}; \tag{3.6}$$

$$f_e \leq C_e, \forall e \in E_f \setminus E_f^{intra}. \tag{3.7}$$

In this formulation, E_f^{intra} is the set of edges between $v_j^{in}, v_j^{out} \in V_B$, $j \in \{1, \cdots, N\}$.

$a_e = A_j$ for $e \in E_f^{intra}$ (corresponding to BS j) and $a_e = 0$ for all the other edges. $z_e = 0$ indicates that the corresponding BS is turned off. The objective (3.5) is to minimize total power consumption based on the general power consumption model with non-zero idle power. Unlike the power-proportional case, the flow problem presented above is known to be the Fixed Charge Network Flow (FCNF) problem [64] that has been shown to be NP-hard. So we can only have polynomial-time heuristic algorithms that give suboptimal solutions. We present an algorithm for the general load allocation problem in the following.

Algorithm 21: The Bilinear Relaxation Based Algorithm

Step 1 Construct the auxiliary flow graph G_f.

Step 2 Solve the problem specified by the MILP-Flow using the bilinear relaxation based algorithm in [64].

Step 3 **forall** $e = (u, v_j^{in})$, where $u = (i, k) \in V_O$ and $v_j^{in} \in V_B$
  $L_{ik}^{jh} = f_e$, where h is the channel assigned
  to BS j;
endforall
return $\langle L_{ik}^{jh} \rangle$.

To the best of our knowledge, the bilinear relaxation based algorithm presented in [64] is the best algorithm for the FCNF problem. The basic idea of this algorithm is to approximate the objective function of the FCNF problem by a piecewise linear one, and construct a Concave Piecewise Linear Network Flow (CPLNF) problem (which can be formulated as an LP problem and solved in polynomial time). A proper choice of parameters in the CPLNF problem can guarantee the equivalence between these two problems. Solving the FCNF problem needs to solve a sequence of CPLNF problems. The algorithm in [64] employs the the bilinear relaxation based algorithm presented in [63], to find a solution to a CPLNF problem. More details can be found in [63] and [64].

We also present a simple algorithm, the *iterative shutdown* algorithm, to solve the load allocation problem without constructing the auxiliary flow graph.

LP-Load(R)

Algorithm 22: The Iterative Shutdown Algorithm

Step 1 $R = \{1, \cdots, N\}$;
Set $l_i := \sum_{k=1}^{K} L_{ik}, \forall i \in \{1, \cdots, N\}$;

Step 2 **while** (1)
$j_{\min} := \text{argmin}_{j \in R} \frac{l_j}{|S_j \bigcap R|}$;
$R := R - \{j_{\min}\}$;
Solve LP-Load(R);
if (No feasible solution) **or** (total power increases)
$R := R + \{j_{\min}\}$;
break;
endif
Update $l_i := \sum_{k=1}^{K} \sum_{j: i \in S_j \bigcap R} l_{jk}^{i}$,
$\forall i \in \{1, \cdots, N\}$ where $\langle l_{jk}^{i} \rangle$ is
the solution returned by solving the LP-Load;
endwhile

Step 3 **return** R and $\langle l_{ik}^{j} \rangle$.

$$\min_{\langle l_{ik}^{j} \rangle} \sum_{i \in R} (A_i + B_i (\sum_{k=1}^{K} \sum_{j: i \in S_j \bigcap R} l_{jk}^{i})) \tag{3.8}$$

Subject to:

$$\sum_{j \in S_i \bigcap R} l_{ik}^{j} = L_{ik}, \ \forall i \in \{1, \cdots, N\}, \\ \forall k \in \{1, \cdots, K\}; \tag{3.9}$$

$$\sum_{k=1}^{K} \sum_{j: i \in S_j \bigcap R} l_{jk}^{i} \leq C_i, \ \forall i \in \{1, \cdots, N\}. \tag{3.10}$$

In the algorithm, R and l_i keep track of the set of active BSs and the load of each BS i respectively. The algorithm keeps trying to turning off a BS until not possible. Every time, a BS j_{min} with smallest load-to-neighbor-number-ratio is selected and the LP-Load is used to test if a feasible load allocation can still be found by shutting down BS j_{min}. The algorithm makes such a selection because it is likely that such a BS can be shut down and its load can be migrated to other active BSs in its neighborhood. The LP-Load is similar to the MILP-Green except that it does not include channel assignment variables y_i^h and the related constraints. Here the superscript h is removed from load allocation variables $\langle i_{jk}^{ih} \rangle$ since the channel assignment is given as input for this algorithm.

4.4 Joint Algorithm

In this section, we present an effective algorithm to jointly solve the channel assignment and load allocation problems. The basic idea of this algorithm is to deal with BS one by one in the descending order of their traffic demands and every time, try to find a feasible load allocation and channel assignment that can lead to minimum power consumption without changing the existing decisions. The algorithm is formally presented as follows.

LP-Load-Local $(i, R, \langle \hat{C}_j \rangle)$

$$\min_{\langle l_{ik}^j \rangle} \sum_{j \in S_i \bigcap R} (A_j + B_j \sum_{k=1}^{K} l_{ik}^j) \tag{4.1}$$

Subject to:

$$\sum_{j \in S_i \bigcap R} l_{ik}^j = L_{ik}, \forall k \in \{1, \cdots, K\}; \tag{4.2}$$

$$\sum_{k=1}^{K} l_{ik}^j \leq \hat{C}_j, \quad \forall j \in S_i \bigcap R. \tag{4.3}$$

In this algorithm, R keeps track of the set of BSs which has been determined (by the algorithm) to power on. $\hat{C}_j$ gives the residual capacity of BS j. $\langle l_{ik}^j \rangle$ specify load allocation and superscript h is removed since the channel assignment of BS in R have already been determined. The algorithm goes through the BS list in the descending of their demands. Every time, the algorithm deals with only one BS. It first check if all its traffic loads can be migrated to BSs in $S_i \bigcap R$ by solving the LP-Load-Local. If so, BS i will be turned off and its load will be migrated to its active neighboring BSs. The algorithm examines this

Algorithm 23: The Joint Algorithm

Step 1 Sort all the BSs in the descending order of its loads and store the sorted list in Q;
$R := \emptyset$;
Set all $\langle \hat{C}_j \rangle := 0$;
Set all $\langle y_i^h \rangle := 0$;

Step 2 **forall** $i \in Q$
  Solve LP-Load-Local($i, R, \langle \hat{C}_j \rangle$)
  if (Found a feasible solution $\langle l_{ik}^j \rangle$)
    $\hat{C}_j := \hat{C}_j - \sum_{k=1}^{K} l_{ik}^j, \forall j \in S_i \bigcap R$;
    continue;
  endif
  $R := R + \{i\}$;
  $h_{\min} = \operatorname{argmin}_{h \in H_i}$LP-Load-Local($i, R, \langle \hat{C}_j \rangle$),
  where $\langle \hat{C}_j \rangle := \langle C_j^h \rangle$;
  if (None of LP-Load-Local returns a
    feasible solution)
    return FAILED;
  endif
  $y_i^{h_{\min}} := 1$;
  $H_j := H_j - \{h_{\min}\}, \forall j \in I_i$;
  $\hat{C}_j := \hat{C}_j - \sum_{k=1}^{K} l_{ik}^j, \forall j \in S_i \bigcap R$;
endforall

Step 3 **return** $\langle y_i^h \rangle$, R, $\langle l_{ik}^j \rangle$.

option first because the idle power usually contributes significantly to the total power usage of a BS (over 50% in most cases), therefore, it is desirable to shut it down if at all possible. Otherwise, the algorithm turns on this BS and finds a channel assignment that can lead to minimum power consumption as well as a load allocation using the same LP. This is a greedy algorithm which makes on/off, channel assignment, and load allocation decisions for a BS in one iteration and will not change them in the following iterations.

The first step of the algorithm can be done in $O(N \log N)$ time. In the 2nd step, even though the LP-Load-Local needs to be solved $O(NH_{\max})$ times, where $H_{\max}$ is the maximum number of available channels on a BS, the LP only includes a small number of variables and constraints since this LP only involves the neighbors of the BS in question. Hence, it can be solved very efficiently.

4.5 Simulation Results

In the simulation, the target region was chosen as a square area with a size of 15×15 km^2. N BSs were randomly placed in the region, whose locations follow a two-dimensional uniform distribution. The effective/maximum transmission ranges were set to $r = 2$km and $r_{\max} = 2r = 4$km, which are quite typical in a cellular network [69]. Similar as in [69], the neighbor set of BS i was defined as $S_i = \{j : d(j,i) \leq (r_{max} + r), j \in \{1, \cdots, N\}\}$. We used a relatively conservative method to define the interference set I_i of BS i as $I_i = \{j : d(j,i) \leq 2r_{max}, j \in \{1, \cdots, N\}\} - \{i\}$.

The total number of available channels was set to 50 and each channel has a bandwidth 50MHz. They were evenly distributed in a portion of spectrum centered at 2GHz. In the simulation, 15 Primary Users (PUs) were randomly placed in the target area. Each PU randomly chose one channel to use. If the distance between a PU and a CR BS is less than the interference range $2r_{\max} = 8$km, then the CR BS cannot be assigned the channel used by the PU. As described above, the capacity of a BS depends on the channel assigned to it. Similar as in [11], the capacity of a BS using the channel at 2GHz was set to 50Mbps. We used the widely-used free-space propagation model [70] and the Shannon's theorem to derive the BS capacities on other channels. To avoid overloading caused by serving migrated traffic loads, we conservatively set the capacity of each BS to 80% of the calculated value.

We used Equation (2.1) to calculate power consumption of a BS. We followed the power consumption settings in [69]. Specifically, we set $A = 2100$W for each BS. The load coefficient B^f was set to 6 for the channel at $f = 2$GHz [69]. As described above this coefficient is channel dependent. B^f consists of two parts: the frequency dependent part $\hat{B}^f$ and the frequency independent part. The value of the frequency dependent part at frequency f can be calculated using the following equation:

$$\hat{B}^f = B^{2G}\gamma + B^{2G}(1-\gamma)\frac{\mu^{2G}}{\mu^f}, \tag{5.1}$$

where μ^f is the efficiency of the material at frequency f and γ is the percentage of channel dependent part, which was set to 0.51 according to [73]. We assumed that LMR-200 [58] is the material used for antenna and transmission line. Then μ^f can be obtained by using the attenuation and power handling calculator [8]. Using this equation, we can find out the values of B^f on different channels.

In all simulation scenarios, the total power consumption of BSs was used as the performance metric and the algorithm presented in Section 4.3.1 was used for determining channel assignment. We used the Gurobi Optimizer 5.0 [36] to solve the MILP-Green to obtain optimal solutions (labeled as "Optimal"). Moreover, we compared our algorithm against a baseline algorithm which uses our algorithm for channel assignment but does not migrate loads or consolidate BSs. All the results presented in the following figures are average over 10 runs and in each run, a different seed was used for random generation of traffic loads.

In this first two scenarios, we evaluated the performance of the min-cost-flow-based algorithm (labeled as "Min-Cost-Flow") for the power-proportional case ($A_i = 0$). In scenario 1, we fixed the number of MVNO $K = 3$ and changed the number of BSs N from 30 to 80 with a step size of 10. In scenario 2, we fixed the number of BSs $N = 50$ and changed the number of MVNOs K from 1 to 6. In these two scenarios, the load of an MVNO on a BS followed a Gaussian distribution with a mean of 5Mbps and a variance of 1Mbps. The simulation results were presented in Figs. 4-2–4-3.

We can make the following observations from these results:

1) The proposed algorithm consistently outperforms the baseline algorithm. On average, it achieves 10% power savings. This shows that even for the power proportional case, migrating traffic loads to more power-efficient BSs can save power. Moreover, the proposed

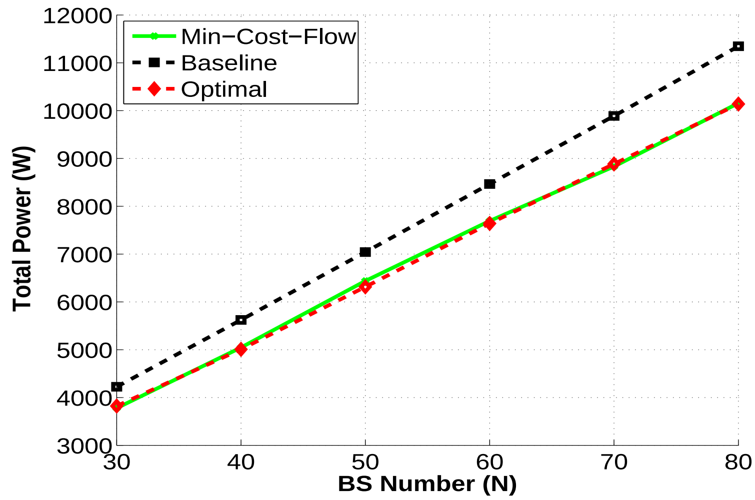


Figure 4-2: Scenario 1: varying BS number (N), power-proportional

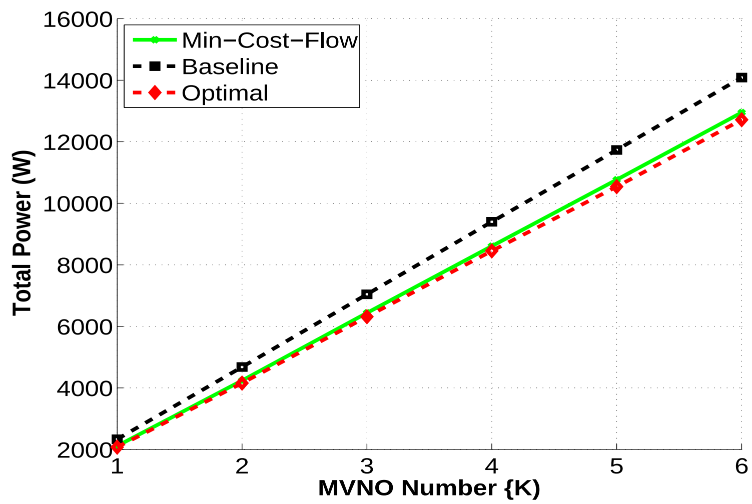


Figure 4-3: Scenario 2: varying MVNO number (K), power-proportional

algorithm produces close-to-optimal solutions since it spends only 1% more power than the optimal on average.

2) No matter which algorithm was used, the power consumption increases linearly with the number of BSs and the number of MVNOs because of the linear power consumption model with $A_i = 0$.

In the next three scenarios, we evaluated the performance of the bilinear relaxation based algorithm (labeled as "Bilinear"), the iterative shutdown algorithm (labeled as "Iterative") and the joint algorithm (labeled as "Joint"). The simulation settings of scenarios 3 and 4 were the same as those of scenarios 1 and 2 respectively. We had an additional scenario, scenario 5, in which we fixed $N = 40$ and $K = 6$, and we changed the load of each BS from $10\% * l$ to $100\% * l$ with a step size of $10\% * l$, where l is the MVNO load on a BS generated using the method described above. The simulation results were presented in Figs. 4-4–4-6.

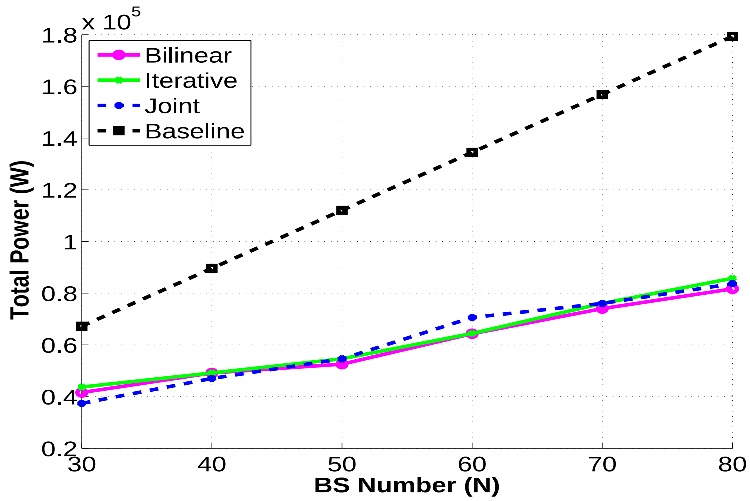


Figure 4-4: Scenario 3: varying BS number (N), the general case

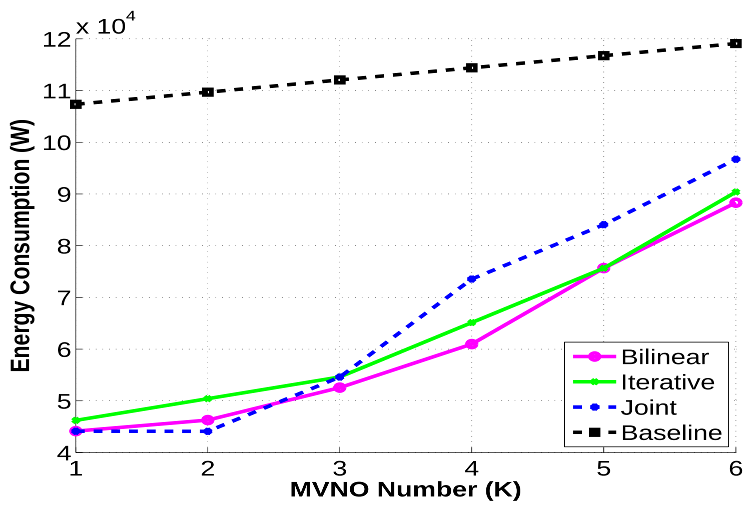


Figure 4-5: Scenario 4: varying MVNO number (K), the general case

We can make the following observations from these results:

1) All the proposed algorithms perform similarly and they all significantly outperform the baseline algorithm. On average, the bilinear relaxation based algorithm, the iterative shutdown algorithm and the joint algorithm, achieve 47%, 46% and 45% power savings

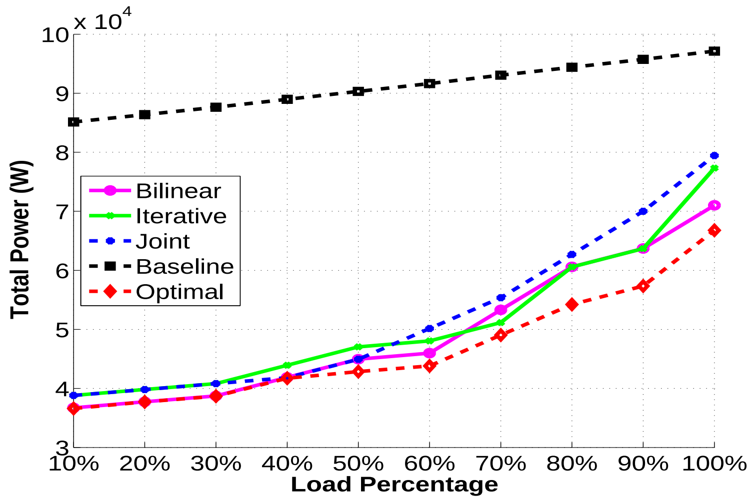


Figure 4-6: Scenario 5: varying MVNO loads, the general case

respectively, which are more significant than those in the power-proportional case. This show that BS consolidation via load migrations can save power significantly, which well justifies the motivation of this work.

2) From Fig. 4-6, we can see all the proposed algorithms yield close-to-optimal performance. Specifically, they spend only 4% more power than the optimal.

3) By increasing the number of BS, the number of MVNOs or MVNO loads (directly), we essentially increase the loads in the network. As expected, the total power consumption increases with the loads no matter which algorithm was used. However, an interesting observation is that the total power consumption does not increases linearly with the loads as what we observed in the power-proportional case. This shows that in the general case, significant power savings come from shutting down BSs because idle power contributes a significant portion of power consumption.

Chapter 5

Conclusions

5.1 Conclusions

In this thesis, we first envisioned the emerging trend of wireless network evolution from 3G/4G to 5G and discussed the requirements and prospects of next generation wireless communications. We introduced some of promising key technologies in 5G and then explained the importance of addressing energy-efficiency. We presented our research works that focused on enabling a greener wireless network by embracing some key promising technologies in 5G. We considered various network design scenarios including both offline planning and online resource allocation. We proposed effective solutions to save power consumption for both wireless infrastucture and UEs.

In Chapter 2, we first studied the *SNR*-Aware Green (SAG) relay placement problem, which sought the multi-hop relay node placement with channel capacity and *SNR* constraints in wireless relay networks. This problem was further divided into two sub-problems, Lower-tier Coverage Relay Allocation (LCRA) problem and Upper-tier Connectivity Relay Allocation (UCRA) problem. For LCRA problem, we provided two approximation algorithms, SAMC and PRO, to solve the problem in two phases. Similarly, for UCRA problem, we proposed a solution consisting of a minimum spanning tree based approximation algorithm

MBMC and an optimal power optimization algorithm UCPO. With solutions to the lower-tier and upper-tier, we combined these solutions of the LCRA and UCRA and presented a solution framework of SAG. Extensive numerical results have been conducted to support our theoretical analysis and showed good performances of our solutions.

In the second part of Chapter 2, we proposed a new perspective on Green Wireless Networking. Without turning off BSs, our proposal was to deploy small cells to offload traffic from macrocell BSs to achieve power saving. We proposed the GREAN problem to find a joint solution for RS placement and RS/BS power consumption to save total network energy. An approximation algorithm was presented for a special case problem GREAG. Then a $(1+\alpha)$-approximation algorithm was presented for the GREAN problem. This work also studied the MC-BAPS problem for multi-cell scenarios. Extensive numerical results have been conducted to support our theoretical analysis and showed that our schemes can provide up to 52% network power consumption compared to traditional wireless macro cell networks.

Besides the offline network planning, we also studied the online network design optimization problems in this thesis. In Chapter 3, we studied green D2D communications in OFDMA-based wireless networks. We formally defined an optimization problem based on a practical model of link data rate, whose objective was to minimize total power consumption while ensuring link data rate requirements. We then presented two joint mode selection, power assignment and channel allocation algorithms, which both solved the problem effectively in polynomial time. Via extensive simulation results, we showed that the proposed algorithms can achieve over 68% power savings, compared to several baseline methods.

In Chapter 4, we considered a power-efficient network planning problem in virtualized CRNs. First, we presented an MILP to provide optimal solutions. Then we presented a general optimization framework to guide algorithm design, which solved two subproblems, channel assignment and load allocation, in sequence. We presented a channel assignment algorithm with an approximation ratio of $(\frac{1}{\Delta})$. For load allocation, we presented a polynomial-time optimal algorithm for the power-proportional case as well as two effective heuristic algorithms for the general case. In addition, we presented a heuristic algorithm that jointly solved the two subproblems. It has been shown by simulation results that the proposed algorithms produced close-to-optimal solutions, and moreover, achieved over 10% and 45% power savings in the power proportional and general cases respectively, compared to a baseline algorithm that did not migrate loads or consolidate BSs.

5.2 Future Research Directions

In this section, we point out future research directions:

- **Resource allocation in heterogenous cellular networks**

  In this thesis, we discussed the base/relay station placement in heterogenous cellular networks, which is a part of offline network planning. However, in online resource allocation, effective and efficient resource allocation schemes need to be developed. Particularly, the resource allocation and interference management in heterogenous networks are even more complicated since both small cells and macrocells share the common cellular spectrum that is very limited to every mobile network operator. Thus, how to address the efficiency of resource allocation in heterogenous networks becomes significant and is worth study.

- **D2D relay in cooperative communications and networking**

  In addition to enabling direct proximal communications, there is also an interesting use case: D2D relay, where a device with better geometry to the BS acts as a relay for another nearby device. A large number of devices, including those in the sleep state, can potentially act as relays and therefore exploit multiuser shadow-diversity. D2D relay introduces new technical problems to be solved, such as discovery of candidate relays, opportunistic relay selection, interference management, multiplexing between access and backhaul links, and minimization of relay power consumption [12]. D2D relay is in essence a special example of two-hop communication, and the concept can be extended to enable multihop and more advanced cooperative communication and networking.

- **Resource scheduling in wireless network virtualization**

  Virtualization has emerged as a useful technology for improving resource utilization and power efficiency. However, research on wireless resource virtualization is still in its infancy. In the last part of this thesis, we studied a power-efficient network planning problem in virtualized cognitive radio networks where multiple mobile virtual network operators (MVNOs) can be supported over a shared physical wireless infrastructure. Each BS can provide service to multiple MVNOs simultanously and meet their traffic demands separately. However, how to schedule radio resource per BS to MVNOs with different demands and how to effectively distribute resource to MVNOs across a field of multiple BSs need to be carefully studied and designed. It is a significant research direction since wireless network virtualization is promised to take place in the near future.

Bibliography

[1] 3GPP TR 32.826. *Telecommunication management; Study on Energy Savings Management (ESM), (Release 10)*. `http://www.3gpp.org/ftp/Specs/html=info/32826.htm`, Mar. 2010.

[2] Ahuja, R. K., Magnanti, T. L., and Orlin, J. B. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, 1993.

[3] Akkarajitsakul, K., Phunchongharn, P., Hossain, E., and Bhargava, V. K. Mode selection for energy-efficient d2d communications in lte-advanced networks: a coalitional game approach. In *Proceedings of IEEE ICCS* (Singapore, Nov. 2012), pp. 488–492.

[4] Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., and Mohanty, S. Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Computer Networks Journal 50*, 13 (2007), 2127–2159.

[5] Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., and Mohanty, S. Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Computer Networks Journal 50*, 13 (2007), 2127–2159.

[6] Alamouti, S. M., and Sharafat, A. R. Resource allocation for energy-efficient device-to-device communication in 4g networks. In *Proceedings of IEEE IST* (Santorini, Greece, Oct. 2014), pp. 1058–1063.

[7] Arnold, O., Richter, F., Fettweis, G., and Blume, O. Power consumption modeling of different base station types in heterogeneous cellular networks. In *IEEE Future Network and Mobile Summit* (2010), pp. 1–8.

[8] The attenuation and power handling calculator. `http://www.timesmicrowave.com/cgi-bin/calculate.pl`.

[9] Bhanage, G., Seskar, I., Mahindra, R., and Raychaudhuri, D. Virtual basestation: architecture for an open shared wimax framework. In *Proceedings of ACM VISA* (New Delhi, India, Sept. 2010), pp. 1–8.

[10] Bhaumik, S., Narlikar, G., Chattopadhyay, S., and Kanugovi, S. Breathe to stay cool: adjusting cell sizes to reduce energy consumption. In *Proceedings of ACM Green Networking* (New Delhi, India, Aug. 2010), pp. 41–46.

[11] Bhaumik, S., Narlikar, G., Chattopadhyay, S., and Kanugovi, S. Breatheto stay cool: adjusting cell sizes to reduce energy consumption. In *Proceedings of ACM SIGCOMM workshop on Green Networking* (New Delhi, India, Aug. 2010), pp. 41–46.

[12] Bhushan, N., Li, J., Malladi, D., Gilmore, R., Brenner, D., Damnjanovic, A., Sukhavasi, R. T., Patel, C., and Geirhofer, S. Network densification: the dominant theme for wireless evolution into 5g. *IEEE Communications Magazine 52*, 2 (2014), 82–89.

[13] Bisio, I., Garibbo, A., Giachetti, S., Marchese, M., Mursia, A., and Raviola, A. Quality of service routing in telecommunications networks: the industrial approach. In *Proceedings of AIRO conference* (2003), pp. 2–5.

[14] Boccardi, F., Heath Jr., R. W., Lozano, A., Marzetta, T. L., and Popovski, P. Five disruptive technology directions for 5g. *IEEE Communications Magazine 52*, 2 (2014), 74–80.

[15] Cognitive radio and cooperative strategies for power saving in multi-standard wireless devices (c2power). `http://www.ict-c2power.eu/`.

[16] Chiang, M., Hande, P., Lan, T., and Tan, C. W. Power control in wireless cellular networks. *Foundations and Trends in Networking 2*, 4 (2008), 381–533.

[17] Chien, C. P., Chen, Y. C., and Hsieh, H. Y. Exploiting spatial reuse gain through joint mode selection and resource allocation for underlay device-to-device communications. In *Proceedings of IEEE WPMC* (Taipel, Taiwan, Sept. 2012), pp. 80–84.

[18] Cover, T., and Gamal, A. E. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory 25*, 5 (1979), 572–584.

[19] Deruyck, M., Tanghe, E., Joseph, W., and Martens, L. Modelling and optimization of power consumption in wireless access networks. *Computer Communications 34*, 17 (2011), 2036–2046.

[20] Deruyck, M., Vereecken, W., Joseph, W., Lannoo, B., Pickavet, M., and Martens, L. Reducing the power consumption in wireless access networks: Overview and recommendations. *Progress In Electromagnetics Research 132* (2012), 255–274.

[21] Doppler, K., Rinne, M., Wijting, C., Ribeiro, C. B., and Hugl, K. Device-to-device communication as an underlay to lte-advanced networks. *IEEE Communications Magazine 47*, 12 (2009), 42–49.

[22] Doppler, K., Yu, C.-H., Ribeiro, C. B., and Janis, P. Mode selection for device-to-device communication underlaying an lte-advanced network. In *Proceedings of IEEE WCNC* (Sydney, Australia, Apr. 2010), pp. 1–6.

[23] Dufková, K., Bjelica, M., Moon, B., Kencl, L., and Boudec, J.-Y. L. Energy savings for cellular network with evaluation of impact on data traffic performance. In

Proceedings of European Wireless Conference (EW) (Lucca, Italy, Apr. 2010), pp. 916–923.

[24] Energy aware radio and network technologies (earth). `http://www.ict-earth.eu/`.

[25] Elayoubi, S. E., Saker, L., and Chahed, T. Optimal control for basestation sleep mode in energy efficient radio access networks. In *Proceedings of IEEE Interference Conference on Computer Communications (INFOCOM)* (Shanghai, China, Apr. 2011), pp. 106–110.

[26] Elgendy, O., Ismail, M., and Elsayed, K. On the relay placement problem in a multi-cell lte-advanced system with co-channel interference. In *Proceedings of IEEE WiMob* (Barcelona, Spain, Oct. 2012), pp. 300–307.

[27] Ericsson AB. *Ericsson mobility report - on the pulse of the networked society*, Nov. 2013. White Paper.

[28] *Flexenclosure E-site Case Study*. `http://www.flexenclosure.com/Esite.aspx`.

[29] *Free-space path loss model*. `http://en.wikipedia.org/wiki/Free-space_path_loss`.

[30] Fusco, G., Buddhikot, M., Gupta, H., and Venkatesan, S. Finding green spots and turning the spectrum dial: Novel techniques for green mobile wireless networks. In *Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)* (Aachen, Germany, May 2011), pp. 613–617.

[31] Gao, C., Sheng, X., Tang, J., Zhang, W., Zou, S., and Guizani, M. Joint mode selection, channel allocation and power assignment for green device-to-device communications. In *Proceedings of IEEE International Conference on Communications (ICC)* (Sydney, Australia, Jun. 2014), pp. 178–183.

[32] Gao, C., Tang, J., Sheng, X., Zhang, W., and Wang, C. Signal-aware green wireless relay network design. In *Proceedings of IEEE International Conference on*

Distributed Computing Systems (ICDCS) (Philadelphia, PA, USA, July 2013), pp. 440–449.

[33] Genc, V., Murphy, S., Yu, Y., and Murphy, J. Ieee 802.16j relay-based wireless access networks: an overview. *IEEE Wireless Communications 15*, 5 (2008), 56–63.

[34] Gesi. `http://www.gesi.org`.

[35] Guo, W., Huang, X., Lou, W., and Liang, C. On relay node placement and assignment for two-tiered wireless networks. *Mobilel Networks and Applications 13*, 1-2 (2008), 186–197.

[36] *Gurobi Optimization*. `http://www.gurobi.com/`.

[37] Hakola, S., Chen, T., Lehtomaki, J., and Koskela, T. Device-to-device (d2d) communication in cellular network-performance analysis of optimum and practical communication mode selection. In *Proceedings of IEEE WCNC* (Sydney, Australia, Apr. 2010), pp. 1–6.

[38] Han, M.-H., Kim, B.-G., and Lee, J.-W. Subchannel and transmission mode scheduling for d2d communication in ofdma networks. In *Proceedings of IEEE VTC Fall* (Québec City, Canada, Sept. 2012), pp. 1–5.

[39] Hasan, Z., Boostanimehr, H., and V.K.Bhargava. Green cellular networks: a survey, some research issues and challenges. *IEEE Communications Surveys and Tutorials 13*, 4 (2011), 524–540.

[40] Hou, Y. T., Shi, Y., Sherali, H. D., and Midkiff, S. F. Prolonging sensor network lifetime with energy provisioning and relay node placement. In *Proceedings of IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)* (Santa Clara, California, USA, Sept. 2005), pp. 295–304.

[41] Huawei Communicate. *The green CDMA base station, Issue 45, 2009*. `http://www.huawei.com/ilink/en/download/HW082746`.

[42] I, C.-L., Rowell, C., Han, S., Xu, Z., Li, G., and Pan, Z. Toward green and soft: a 5g perspective. *IEEE Communications Magazine 52*, 2 (2014), 66–73.

[43] IEEE 802.16E Working Group. *Part16: Air interface for fixed and mobile broadband wireless access systems amendment 2: Physical and Medium Access Control layers for combined fixed and mobile operation in licensed bands*, 2005. IEEE Standard.

[44] Islam, M. M., Hassan, M. M., Lee, G.-W., and Huh, E.-N. A survey on virtualization of wireless sensor networks. *IEEE Sensors Journal 12*, 2 (2012), 2175–2207.

[45] ITU-T Focus Group on Future Networks (FG FN), FG-FN OD-66. *Draft Deliverable on "Overview of Energy Saving of Networks"*. `http://www.itu.int/dmspub/itut/oth/3A/05/T3A050000660001MSWE.doc`, Oct. 2010.

[46] Jung, M., Hwang, K., and Choi, S. Joint mode selection and power allocation scheme for power-efficient device-to-device (d2d) communication. In *Proceedings of IEEE VTC Spring* (Yokohama, Japan, May 2012), pp. 1–5.

[47] Khakurel, S., Mehta, M., and Karandikar, A. Optimal relay placement for coverage extension in lte-a cellular systems. In *Proceedings of National Conference on Communications* (Kharagpur, India, Feb. 2012), pp. 1–5.

[48] Kokku, R., Mahindra, R., Zhang, H., and Rangarajan, S. Nvs: a substrate for virtualizing wireless resources in cellular networks. *IEEE/ACM Transactions on Networking 20*, 5 (2012), 1333–1346.

[49] Kuri, T., Harai, H., Wada, N., Kawanishi, T., and Hosokawa, M. Adaptable access system: pursuit of ideal future access system architecture. *IEEE Network Magazine 26*, 2 (2012), 42–48.

[50] López-Pérez, David, A. L., Jüttner, A., Rivano, H., and Zhang, J. Optimization method for the joint allocation of modulation schemes, coding rates, resource blocks and power in self-organizing lte networks. In *Proceedings of IEEE International Conference on Computer Communications (INFOCOM)* (Shanghai, China, Apr. 2011), pp. 111–115.

[51] Lei, L., Zhong, Z., Lin, C., and Shen, X. Operator controlled device-to-device communications in lte-advanced networks. *IEEE Wireless Communications 19*, 3 (2012), 96–104.

[52] Li, L. E., Mao, Z. M., and Rexford, J. Cellsdn: Software-defined cellular networks. Tech. rep., Princeton University, 2012.

[53] Li, Y., Jin, D., Gao, F., and Zeng, L. Joint optimization for resource allocation and mode selection in device-to-device communication underlaying cellular networks. In *Proceedings of IEEE ICC* (Sydney, Australia, Jun 2014), pp. 2245–2250.

[54] Lin, B., Ho, P.-H., Xie, L.-L., and Shen, X. Optimal relay station placement in ieee 802.16j networks. In *Proceedings of ACM International Conference on Wireless Communications and Mobile Computing (IWCMC)* (Honolulu, Hawaii, Aug. 2007), pp. 25–30.

[55] Lin, B., Ho, P.-H., Xie, L.-L., and Shen, X. S. Relay station placement in ieee 802.16 j dual-relay mmr networks. In *Proceedings of IEEE International Conference on Communications (ICC)* (Beijing, China, May 2008), pp. 3437–3441.

[56] Lin, G.-H., and Xue, G. Steiner tree problem with minimum number of steiner points and bounded edge-length. *Information Processing Letters 69*, 2 (1999), 53–57.

[57] Liu, Z., Peng, T., Xiang, S., and Wang, W. Mode selection for device-to-device (d2d) communication under lte-advanced networks. In *Proceedings of IEEE ICC* (Ottawa, Canada, Jun 2012), pp. 5563–5567.

[58] Lmr-200. `http://www.exergia.info/books/cables/LMR200A.pdf`.

[59] Lorincz, J., Garma, T., and Petrovic, G. Measurementsandmodellingofbase station power consumption under real traffic loads. *Sensors 12*, 4 (2012), 4181–4310.

[60] Marsan, M. A., Chiaraviglio, L., Ciullo, D., and Meo, M. Optimal energy savings in cellular access networks. In *Proceedings of IEEE Interference Conference on Communications (ICC) GreenComm Workshop* (Dresden, Germany, Jun. 2009), pp. 1–5.

[61] Micallef, G., Mogensen, P., and Scheck, H. O. Cell size breathing and possibilities to introduce cell sleep mode. In *Proceedings of European Wireless Conference (EW)* (Lucca, Italy, Apr. 2010), pp. 111–115.

[62] Mustafa, N., and Ray, S. Ptas for geometric hitting set problems via local search. In *Proceedings of ACM annual symposium on Computational geometry* (Aarhus, Denmark, Jun. 2009), pp. 17–22.

[63] Nahapetyan, A., and Pardalos, P. A bilinear relaxation based algorithm for concave piecewise linear network flow problems. *Journal of Industrial and Management Optimization 3*, 1 (2007), 71.

[64] Nahapetyan, A., and Pardalos, P. Adaptive dynamic cost updating procedure for solving fixed charge network flow problems. *Computational Optimization and Application 39*, 1 (2008), 37–50.

[65] Nakauchi, K., Ishizu, K., Murakami, H., Nakao, A., and Harada, H. Amphibia: a cognitive virtualization platform for end-to-end slicing. In *Proceedings of IEEE ICC* (Kyoto, Japan, Jun. 2011), pp. 1–5.

[66] Nokia Siemens Networks. *Multiradio Base Station makes network evolution easier and greener than ever, Press Release Feb 5 2009.* `http://www.nokiasiemensnetworks.com/sites/default/files/document/NokiaSiemensNetworks20090205enFlexiMultiradioBTS.pdf`.

[67] Pan, J., Hou, Y. T., Cai, L., Shi, Y., and Shen, S. X. Topology control for wireless mobile networks. In *Proceedings of ACM International Conference on Mobile Computing and Networking (MobiCom)* (San Diego, California, USA, Sept. 2003), pp. 286–299.

[68] Pei, Y., and Mutka, M. Joint bandwidth-aware relay placement and routing in heterogeneous wireless networks. In *Proceedings of IEEE ICPADS* (Tainan, Taiwan, Dec. 2011), pp. 420–427.

[69] Peng, C., Lee, S.-B., Lu, S., Luo, H., and Li, H. Traffic-driven power saving in operational 3g cellular networks. In *Proceedings of ACM International Conference on Mobile Computing and Networking (MobiCom)* (Las Vegas, Nevada, USA, Sept. 2011), pp. 121–132.

[70] Rappaport, T. S. *Wireless Communications: Principles and Practice*, 2nd ed. Prentice-Hall, 2002.

[71] Sakai, S., Togasaki, M., and Yamazaki, K. A note on greedy algorithms for the maximum weighted independent set problem. *Discrete Applied Mathematics 126*, 2 (2003), 313–322.

[72] Saker, L., Elayoubi, S., and Scheck, H. System selection and sleep mode for energy saving in cooperative 2g/3g networks. In *Proceedings of IEEE Vehicular Technology Conference Fall (VTC Fall)* (Anchorage, Alaska, USA, Sept. 2009), pp. 1–5.

[73] Son, K., and Krishnamachari, B. Speedbalance: speed-scaling-aware optimal load balancing for green cellular networks. In *Proceedings of IEEE Interference Conference on Computer Communications (INFOCOM)* (Orlando, Florida, USA, Mar. 2012), pp. 2816–2820.

[74] T. C. Group. *Smart 2020: Enabling the low carbon economy in the information age*, 2008.

[75] Tang, J., Misra, S., and Xue, G. Joint spectrum allocation and scheduling for fair spectrum sharing in cognitive radio wireless networks. *Computer Networks Journal 52*, 11 (2003), 2148–2158.

[76] Tao, Z., Teo, K. H., and Zhang, J. Aggregation and concatenation in ieee 802.16j mobile multihop relay (mmr) networks. In *Proceedings of IEEE Mobile WiMAX Symposium* (Orlando, FL, USA, Mar. 2007), pp. 85–90.

[77] Thompson, J., Ge, X., Wu, H.-C., Irmer, R., Jiang, H., Gerhard, F., and Alamouti, S. 5g wireless communication systems: prospects and challenges [guest editorial]. *IEEE Communications Magazine 52*, 2 (2014), 62–64.

[78] Towards real energy-efficient network design (trend). `http://www.fp7-trend.eu/`.

[79] UMTS Forum. *Mobile traffic forecasts 2010-2020 report*, Jan. 2011. Report No. 44.

[80] Wang, C.-X., Haider, F., Gao, X., You, X.-H., Yang, Y., Yuan, D., Aggoune, H. M., Haas, H., Fletcher, S., and Hepsaydir, E. Cellular architecture and key technologies for 5g wireless communication networks. *IEEE Communications Magazine 52*, 2 (2014), 122–130.

[81] Xia, L., Kumar, S., Yang, X., Gopalakrishnan, P., Liu, Y., Schoenberg, S., and Guo, X. Virtual wifi: bring virtualization from wired to wireless. *ACM SIGPLAN Notices 46*, 7 (2011), 181–192.

[82] Xiang, S., Peng, T., Liu, Z., and Wang, W. A distance-dependent mode selection algorithm in heterogeneous d2d and imt-advanced network. In *Proceedings of IEEE Globecom Workshops* (Anaheim, California, USA, Dec. 2012), pp. 416–420.

[83] Xiang, S., Quan, Q., Peng, T., and Wang, W. Performance analysis of cooperative mode selection in hybrid d2d and imt-advanced network. In *Proceedings of IEEE ChinaCom* (Kunming, China, Aug. 2012), pp. 717–721.

[84] Xu, K., Hassanein, H., Takahara, G., and Wang, Q. Relay node deployment strategies in heterogeneous wireless mobile networks: multiple-hop communication case. In *Proceedings of IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)* (Santa Clara, California, USA, Sept. 2005), pp. 575–585.

[85] Yu, C. H., Doppler, K., Ribeiro, C. B., and Tirkkonen, O. Resource sharing optimization for device-to-device communication underlaying cellular networks. *IEEE Transactions on Wireless Communications 10*, 8 (2011), 2752–2763.

[86] Yu, G., Xu, L., Feng, D., Yin, R., Li, G. Y., and Jiang, Y. Joint mode selection and resource allocation for device-to-device communications. *IEEE Transactions on Communications 62*, 11 (2014), 3814–3824.

[87] Yucek, T., and Arslan, H. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Communications Surveys & Tutorials 11*, 1 (2009), 116–129.

[88] Yun, D., and Yi, Y. Virtual network embedding in wireless multihop networks. In *Proceedings of ACM CFI* (Seoul, Korea, Jun. 2011), pp. 30–33.

[89] Zaki, Y., Zhao, L., Goerg, C., and Timm-Giel, A. Lte mobile network virtualization. *Mobile Networks and Applications 16*, 4 (2011), 424–432.

[90] Zhang, W., Bai, S., Xue, G., Tang, J., and Wang, C. Darp:distance-aware relay placement in wimax mesh networks. In *Proceedings of IEEE International Conference on Computer Communications (INFOCOM)* (Shanghai, China, Apr. 2011), pp. 2060–2068.

[91] Zhou, H., Ji, Y., Li, J., and Zhao, B. Joint mode selection, mcs assignment, resource allocation and power control for d2d communication underlaying cellular networks. In *Proceedings of IEEE WCNC* (Istanbul, Turkey, Apr. 2014), pp. 1667–1672.

[92] Zhu, K., and Hossain, E. Joint mode selection and spectrum partitioning for device-to-device communication: a dynamic stackelberg game. *IEEE Transactions on Wireless Communications 14*, 3 (2015), 1406–1420.

[93] Zhu, Z., Gupta, P., Wang, Q., Kalyanaraman, S., Lin, Y., Franke, H., and Sarangi, S. Virtual base station pool: towards a wireless network cloud for radio access networks. In *Proceedings of ACM International Conference on Computing Frontiers* (Ischia, Italy, May 2011), p. 34.

CHENFEI GAO

206 Fellows Ave. ⋄ Syracuse, NY 13210
(315) · 560 · 7658 ⋄ cgao03@syr.edu

AREAS OF INTEREST

Prototype and Testbed Implementation
Software-Defined Networking
Cloud-Based Wireless Networking
Wireless Network Virtualization

EDUCATION

Syracuse University, Syracuse, NY	Aug 2011 - Oct 2015
Ph.D. candidate in Electrical & Computer Engineering	GPA 3.9/4.0
Syracuse University, Syracuse, NY	Aug 2009 - May 2011
M.S. in Electrical Engineering	GPA 3.83/4.0
Beijing University of Posts and Telecommunications, Beijing, China	Sept 2005 - Jun 2009
B.S. in Electrical Engineering	

PROFESSIONAL SKILL

OpenFlow Tools	Floodlight, Ryu, Beacon; Mininet; Open vSwitch; Wireshark
Cloud Computing Tools	OpenStack, VMware, KVM
Programming Languages	C, C++, Java, Python, JavaScript, Matlab
Programming Techniques	Unix Network Programming, Design Patterns, Shell Scripts
Platforms & APIs	USRP, GNURadio; Gurobi, JSON, ZeroMQ, Gstreamer, LEDA
Databases & OS	MySQL, Oracle; Linux Ubuntu, MacOS, Windows
Optimization Techniques	LP, ILP, QP, CP, Stochastic Programming, Algorithm Design
Other Tools	Git, SVN, Vim, Emacs, Eclipse, Visual Studio, IntelliJ IDEA

RESEARCH & ENGINEERING EXPERIENCE

Syracuse University, Department of Electrical Engineering & Computer Science
Graduate Research Assistant, Computer Networking Laboratory

* **Project: Cloud-based Wireless Virtualization Framework Design and Implementation**
· Designed novel network architecture for cloud-based wireless RAN virtualization model.
· Developed mechanisms in both cloud and BSs to handle buffer overflowed and user association issue.
· Implemented the prototype using *Software-Defined Radios (SDR) such as USRP N200s* based on *GNURadio platform* and established large-scale simulations by writing a *simulator in C++ and Matlab.*

* **Project: Video Transmission Experiment in Software Defined Networks**
· Established a *testbed of Pica8 OpenFlow compatible switches, two video servers and multiple clients.*
· Developed efficient video traffic routing scheme to achieve load balancing and maximum throughput.
· Implemented the prototype by using *Floodlight SDN controller, Open vSwitch and Gstreamer APIs.*
· Wrote python application in controller to configure switches using *Floodlight REST APIs.*
· Wrote the simulations in *Mininet* involving *Floodlight, Wireshark* and *linux traffic generation tools.*

* **Project: Explicit Inter-Data-Center Routing with SDN**
· Designed *OpenFlow-based* achitecture for multipath routing demands in inter-datacenter WANs.
· Focused on developing dependable multipath routing scheme to achieve maximum utility.
· Implemented a *simulator in C++ running on DARPA CORONET 60-node topology* and randomly generated topologies to verify the suitability and computational scalability of proposed schemes.

* **Air Force Research Laboratory Project in Phase I:**
Dynamic Cross-layer Routing Using Cognitive Spectrum Allocation Techniques
· Designed adaptive networking protocols to achieve network resiliency in contested radio spectra.

- Responsible for implementing a *SDR-based testbed* to show joint optimization of network routing and spectrum allocation on OSSIE platform.
- Wrote *OSSIE blocks in C++* to detect energy level on spectra then assign route, channel for transmission .

* **Research Project: Going Direct - Device-to-Device Underlaying Cellular Network Design**
- Designed new D2D communication network architecture to achieve traffic offloads from BSs.
- Researched on an optimization problem based on a practical link data rate model, whose objective is to minimize power consumption while meeting user data rate requests.
- Proposed new algorithms to jointly determine mode selection, channel allocation and power assignment.

* **Research Project: "Relax, but Do Not Sleep"**
- Designed new cellular network mechanism by heterogeneous cell deployment to save power costs.
- Focused on solving the problem of minimizing the total power consumption in RAN by tactically reducing the coverage of each BS and strategically placing microcells to offload the traffic while satisfying the coverage constraints.
- Proposed approximation algorithms to solve the problem in both single-cell and multi-cell scenarios.

* **Research Project: Signal-Aware Relay Station Placement in Green Cellular Network**
- Designed novel two-tiered relay network to enhance network coverage and capacity.
- Focused on solving the problems of signal-aware relay station placement and power allocation.
- Solved optimization problems by mixed integer linear programming and approximation algorithms.

PUBLICATIONS

Greening Wireless Relay Networks: A SNR-Aware Approach
Parallel and Distributed Systems, IEEE Transactions on, (TPDS), accepted
Authors: Chenfei Gao, Jian Tang, Xiang Sheng, Weiyi Zhang, Chonggang Wang

Enabling Green Wireless Networking with Device-to-Device Links: A Joint Optimization Approach
Wireless Communications, IEEE Transactions on, (TWC), accepted with minor revisions
Authors: Chenfei Gao, Jian Tang, Xiang Sheng, Weiyi Zhang, Shihong Zou, Mohsen Guizani

Joint Mode Selection, Channel Allocation and Power Assignment for Green D2D Communications
IEEE International Conference on Communications (**ICC**) 2014, accepted, *Best Paper Award*
Authors: Chenfei Gao, Xiang Sheng, Jian Tang, Weiyi Zhang, Shihong Zou and Mohsen Guizani

Relax, but Do Not Sleep: A New Perspective on Green Wireless Networking
IEEE International Conference on Computer Communications (**INFOCOM**) 2014, accepted
Authors: Chenfei Gao, Weiyi Zhang, Jian Tang, Chonggang Wang, Shihong Zou and Sen Su

Signal-Aware Green Wireless Relay Network Design
International Conference on Distributed Computing Systems (**ICDCS**) 2013, accepted
Authors: Chenfei Gao, Jian Tang, Xiang Sheng, Weiyi Zhang, Chonggang Wang

Leveraging Load Migration and Basestaion Consolidation for Green Communications in Virtualized Cognitive Radio Networks
IEEE International Conference on Computer Communications (**INFOCOM**) 2013, accepted
Authors: Xiang Sheng, Jian Tang, Chenfei Gao, Chonggang Wang, Weiyi Zhang

SOR: An Objective Ranking System Based on Mobile Phone Sensing
International Conference on Distributed Computing Systems (**ICDCS**) 2014, accepted
Authors: Xiang Sheng, Jian Tang, Jing Wang, Chenfei Gao, Guoliang Xue

INTERNSHIP EXPERIENCE

Critical Technologies Inc. Jun 2011 - Mar 2012
Software Developer Intern Syracuse, NY

PROFESSIONAL ACTIVITIES AND AWARDS

Awards and Honors
2014 Best Paper Award in mobile ad hoc symposium of 2014 IEEE ICC
2010 Excellent Graduate Student Award by EECS department in Syracuse University

Journal Reviewer

IEEE Journal on Selected Areas in Communications (JSAC)
IEEE Transactions on Vehicular Technology
KSII Transactions on Internet and Information Systems