Strategic Location Planning for Broadband Access Networks under Cooperative Transmission

by

Bin Lin

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

To achieve a cost-effective network deployment, employing state-of-art technical advances provides a practical and effective way to enhance system performance and quality of service provisioning. Cooperative transmission has been recognized as one of the most effective paradigms to achieve higher system performance in terms of lower bit-error rate, higher throughput, larger coverage, more efficient energy utilization, and higher network reliability. This dissertation studies the location planning for the deployment of broadband access networks and explores the great potential of cooperative transmission in the context of single-cell cooperative relaying and multi-cell cooperative transmission, respectively. The placement problem is investigated in two categories of network deployment environment, i.e., an existing wireless access network and a perspective broadband access network, respectively.

In an existing wireless access network, to solve some practical problems such as the requirements of capacity enhancement and coverage extension, relay stations (RSs) are introduced in the network architecture. We propose two optimization frameworks with the design objectives of maximizing cell capacity and minimizing number of RSs for deployment, respectively. Mathematical formulations are provided to precisely capture the characteristics of the placement problems. The corresponding solution algorithms are developed to obtain the optimal (or near-optimal) results in polynomial time. Numerical analysis and case studies are conducted to validate the performance benefits due to RS placement and the computation efficiency of the proposed algorithms.

To deploy a new metropolitan-area broadband access network, we explore the integration of passive optical network (PON) and wireless cooperative networks (WCN)

under the multi-cell cooperative transmission technology. An optimization framework is provided to solve the problem of dimensioning and site planning. The issues of node placement, BS-user association, wireless bandwidth and power breakdown assignment are jointly considered in a single stage to achieve better performance. We also propose a solution to the complex optimization problem based on decomposition and linear approximation. Numerical analysis and case studies are conducted to verify the proposed framework. The results demonstrate the performance gains and economic benefits.

Given a set of network parameters, the proposed optimization frameworks and solutions proposed in this dissertation can provide design guidelines for practical network deployment and cost estimations. And the constructed broadband access networks show a more cost-effective deployment by taking advantage of the cooperative transmission technology.

Acknowledgments

Numerous people have supported me during the development of this dissertation. A few words mention here cannot adequately capture all my appreciation.

I would like to express my deepest and sincerest gratitude to my supervisor Professor Pin-Han Ho for his support and guidance during my study at the University of Waterloo. He has provided me with a motivating, enthusiastic, and critical research atmosphere. It is my honor to pursue my Ph.D. study under his supervision. He has guided me not only the way to do research but also the attitude in life, which are of great benefit to me forever.

My extreme appreciation goes to my thesis committee members: Professor Liang-Liang Xie, Professor Sagar Naik, Professor Xinzhi Liu, and Professor Chadi M. Assi. They contributed their precious time to read my thesis, and provided valuable suggestions and comments that helped to improve the quality of this thesis.

I am deeply grateful to Professor Xuemin (Sherman) Shen. My research benefits a lot from his solid knowledge and professional view in wireless communications and networking. Many thanks to Professor Weihua Zhuang. From her class, I obtained solid theoretical background.

My special thanks are to Dr. Fen Hou, Dr. Wei Song, Dr. Ping Wang, Ms. Lin Cai, Ms. Xiaojing Meng, Ms. Huyu Wu and Ms. Li Wang for their friendship during the last few years, which makes my life abroad easier and bring me many warm and happy memories.

I also wish to thank Dr. Janos Tapolcai and Dr. Bin Wu for their intelligence and expertise on optimization that particularly contributed to this thesis's research, and Dr. Ching-Yu Liao, Dr. Mehri Mehrjoo, Dr. Xinhua Ling, Dr. Stanley Liu, Dr. Hai

Jiang, Mr. Haojin Zhu, Mr. Yanfei Fan, and other colleagues for their friendly help, stimulating discussions, and valuable suggestions. It is my great pleasure to be in a wonderful research team and a large warm family.

I want to extend my gratitude to the administrative staff at the Department of Electrical and Computer Engineering: Ms. Wendy Boles, Ms. Annette Dietrich, Ms. Lisa Hendel, and Ms. Karen Schooley for their kindness and help.

As always, I thank my family for inspiring me to pursue an academic career: Mom, Dad and sister. Their encouragement, support, kindest help and constructive advices through these many years have meant more I can ever express. I would also like to express my deep gratitude to my forever beloved grandpa and grandma. The memory they had gave me is one of the most cherished in my life.

Last, my most tender and sincere thanks go to my loving husband, *Qiang*, for his support to me in innumerable ways, even from thousands of miles away. I am greatly indebted to him for his love, patience and understanding, which are indispensable for the completion of my Ph.D. program.

To my beloved parents, sister and husband Qiang

Contents

List of Tables **xi**

List of Figures **xii**

List of Abbreviations **xv**

1 Introduction **1**

1.1 Broadband Access Networks 2

1.2 Wireless Cooperative Transmissions 4

1.2.1 Single-cell Cooperative Relaying 4

1.2.2 Multi-cell Cooperative Transmission 5

1.3 Importance and Challenges of Location Planning 6

1.4 Research Motivations, Objectives and Contributions 8

1.5 Outline of the Dissertation 12

2 Background and Literature Review **13**

2.1 Cooperative Transmissions in Wireless Networks 13

2.2 Location Planning in Communication Networks 18

2.3 Integration of Optical and Wireless Networks 23

2.4 Summary 24

3 Capacity Enhancement with Relay Station Placement **26**

3.1 System Model 27

3.1.1 Relay-Based Wireless Network Architecture 27

3.1.2 Cooperative Relaying Strategy 29

3.2 Problem Formulation 30

3.2.1 An Mixed Integer Nonlinear Model 30

3.2.2 An Mixed Integer Linear Model 33

3.2.3 Upper Bound of the Cell Capacity 35

3.3 Solution Algorithms to RS Placement 36

3.3.1 An Improved Genetic Algorithm 36

3.3.2 A Fast Heuristic Algorithm 42

3.4 Numerical Results 44

3.4.1 Case Study I: Single Relay 45

3.4.2 Case Study II: Multiple RSs and SSs 47

3.5 Summary 49

4 Location Planning under Multi-level Cooperative Relaying **55**

4.1 Modeling of Multi-level Cooperative Relaying 56

4.2 Problem Formulation 57

4.3 An Optimal Dimensioning and Location Planning Algorithm 64

4.3.1 Algorithm Description 64

4.3.2 Computational Complexity Analysis 70

4.3.3 Optimality Proof 70

4.4 Numerical Results 74

4.4.1 Case Study I: 1-D Multi-level Cooperative Relaying 74

4.4.2 Case Study II: 2-D Multi-level Cooperative Relaying 75

4.5 Summary 78

5 Location Planning in Integrated PON-WCN 79

5.1 System Model of Integrated PON-WCN 80

5.1.1 Main Entities in Integrated PON-WCN 81

5.1.2 Integrating Mode 82

5.1.3 Cooperative Transmissions in Integrated PON-WCN 83

5.1.4 Radio Propagation and Coverage Model 85

5.1.5 Optical Propagation Model in Fibers 85

5.2 Problem Formulation 86

5.3 Solution 94

5.3.1 Problem Decomposition 94

5.3.2 Linear Approximation Based Reformulation 95

5.4 Numerical Results and Case Studies 99

5.5 Summary 106

6 Conclusions and Future Research 110

6.1 Major Research Contributions 110

6.2 Future Research 111

6.3 Final Remarks 114

A Appendix 115

A.1 Calculation of the Maximal Chord-Curve Distance D_{UB}. 115

Bibliography 117

List of Tables

3.1 Definitions of Important Symbols for Problem CMRP 50

3.2 Problem Size, Computation Time, Amount of Memory Occupied for Cplex Solving the CMRP-MILP Formulation 52

3.3 GA Parameter Setting 52

3.4 Comparison in Computation Time and Optimality Gap against the Benchmark of CPLEX 52

4.1 Definitions of Important Symbols for Problem DLP 58

4.2 Objective Value Comparison 78

5.1 Definitions of Important Symbols for Problem DSP-PW 88

5.2 Problem Size, Average Computation Time and Average Optimality Gap for CPLEX Solving Problem DSP-PW 102

5.3 BU-Association and Spectra Efficiency of VIPs in Scenario (I) for CPLEX Solving Problem DSP-PW 107

List of Figures

1.1 Single-cell cooperative relaying. 5

1.2 Multi-cell cooperative transmission. 6

2.1 The BC/MAC in three-node relay channel model 14

3.1 A relay-based broadband wireless access network architecture. 27

3.2 The three-node relay channel model. 29

3.3 Population and individual (chromosome) representation for GA-based heuristic to Problem CMRP 37

3.4 Contour of rate gain for corresponding SS, the distance between the BS and SS is normalized, $P_{BS} = 1W$, $P_{RS} = 0.5W$, $\alpha = 2$, $G_a = 0.8bit/sec/Hz$. 41

3.5 Achievable rate (in unit of bits/sec/channel use) comparison in 3-node relay model. 45

3.6 2D contour of relative rate gain. 46

3.7 Illustration of the cell layout. 51

3.8 Convergence comparison of GA and improved GA. 53

3.9 Cell capacity vs. number of RSs in Scenario (I). 53

3.10 Achievable rate for each SS where 12 RSs are deployed. 54

4.1 An illustration of 3-level relaying. 56
4.2 An illustration of practical relay routes and virtue flow (BS→ SS_2). . . 60
4.3 The virtue flow matrix $\mathbb{F}$ for SS_2. 61
4.4 The compressed virtue flow matrix $\mathbb{F}$ for SS_2. 61
4.5 An illustration of FRS table ($ML = 3$). 66
4.6 The rectilinear multi-level CR model. 74
4.7 The achievable rate of destination vs. distance between source and destination. 75
4.8 The optimal results obtained by proposed 2-phase algorithm. 76
4.9 The comparison of computation time of proposed algorithm against exhaustive search. 77

5.1 Integrated PON-WCN model 81
5.2 Model decomposition structure of problem (DSP-PW). 95
5.3 Linear approximation of the convex function $f(\xi_k)$ 96
5.4 Flow chart of the ε-approximation algorithm to linearize curve $f(\xi_k)$. . 98
5.5 2-D contour of achievable rate: comparison in CT and NCT mode. . . 100
5.6 Illustration of network layout of AOI in Scenario (I). 101
5.7 The perceived SNR and the number of associated CPs for each TP in Scenario (I). 103
5.8 The number of ONU-BSs vs. the increasing ratio of total traffic demands of all VIPs. 108
5.9 The layout of PSs, CPs, and obtained ONU-BSs after solving Subproblem 1 (DSP-W) in Scenario (I). 109
5.10 The resultant optimal PON tree (including fiber layout and splitter placement) in Scenario (I). 109

A.1 Maximal chord-curve distance. 116

List of Abbreviations

1-D	One Dimensional
2-D	Two Dimensional
3GPP	The Third Generation Partnership Project
A-F	Amplify-and-Forward
AON	Active Optical Networks
AP	Access Point
AWGN	Additive White Gaussian Noise
BAN	Broadband Access Networks
BER	Bit Error Rate
BS	Base Station
BWA	Broadband Wireless Access
CAPEX	Captical Expenditure
CDMA	Code Division Multiple Access
C-F	Compress-and-Forward
CMRP	Capacity Maximization with RS Placement
CO	Central Office
CP	Candidate Position
CR	Cooperative Relaying

CT	Cooperative Transmissions
D-F	Decode-and-Forward
DLP	Dimensioning and Location Planning
DSL	Digital Subscriber Line
DSP	Dimensioning and Site Planning
FRS	Feasible-Relay-Sequence
EPON	Ethernet Passive Optical Networks
FMC	Fixed Mobile Convergence
FTP	File Transfer Protocol
GA	Genetic Algorithm
GPON	Gigabit Passive Optical Networks
HFC	Hybrid Fiber Coax
IEEE	Institute of Electrical and Electronics Engineering
IETF	Internet Engineering Task Force
ILP	Integer Linear Program
IPTV	Internet Protocol TeleVision
LB	Lower Bound
LOS	Line of Sight
LTE	Long Term Evolution
MAC	Media Access Control
MAN	Metropolitan Area Network
MIMO	Multiple Input Multiple Output
MINLP	Mixed Integer Non-Linear Program
MISO	Multiple Input Single Output

MMR	Mobile Multi-hop Reply
NCT	Non-Cooperative Relaying
NLOS	Non Line of Sight
NP-HARD	Nondeterministic Polynomial-time hard
OFDM	Orthogonal Frequency Division Multiplex
OFDMA	Orthogonal Frequency Division Multiple Access
OLT	Optical Line Terminal
ONU	Optical Network Unit
OPEX	Operational Expenditure
PC	Personal Computer
P2P	Peer to Peer
PON	Passive Optical Networks
PS	Potential Site
PHY	Physical
PMP	Point-to-Multi-Point
QoS	Quality of Service
RBA	Resource Breakdown Assignment
RS	Relay Station
ROF	Radio-Over-Fiber
SNR	Signal-to-Noise Ratio
SS	Subscriber Station
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TDD	Time Division Duplexing

TDM	Time Division Multiplexing
TP	Test Point
UB	Upper Bound
UT	User Terminal
UWB	Ultra-WideBand
UMTS	Universal Mobile Telecommunications Systems
VIP	Very Important Point
VoIP	Voice over Internet Protocol
WCN	Wireless Cooperative Network
WDM	Wavelength Division Multiplexing
WiMAX	Worldwide Interoperability Microwave Access
Wi-Fi	Wireless Fidelity
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Networks
WPAN	Wireless Personal Area Networks
WSN	Wireless Sensor Networks

Chapter 1

Introduction

This chapter begins with a brief introduction to broadband access networks and their developing trends in Section 1.1. Due to the importance of incorporating state-of-the-art technologies, wireless cooperative transmission that can provide significant performance benefits has drawn great attention. In Section 1.2, we explore the wireless cooperative transmission technologies in the context of single-cell cooperative relay transmission and multi-cell cooperative transmission, respectively. In particular, we focus on one of the key factors, that is, the locations of the main network equipments involved in cooperative transmission. The importance and research challenges of location planning is introduced in Section 1.3. The real-world deployment demands may trigger a new round of location planning, which is the research focus of this dissertation. We target at proposing cost-effective solutions to optimize the placement by taking advantage of cooperative transmission, which has seldom been considered in previous study of location planning. Section 1.4 introduces the research motivations, objectives and contributions.

1.1 Broadband Access Networks

A Broadband access network (BAN) refers to the portion of a high-speed data communication network that connects to a subscriber's residence, place of business or mobile location [1]. Based on the type of "last-mile" transmission media, BANs can be divided into two categories: one is the wire-line network, such as Digital Subscriber Line (DSL) network, coaxial cable network, and optical fiber network; the other is the wireless network, such as IEEE 802.11-based Wireless Local Area Network (WLANs) or WiFi, IEEE 802.16-based Wireless Metropolitan-Area Network (WMANs) or WiMAX, third-generation (3G) Long Term Evolution (LTE) cellular network.

The wired and wireless networks currently co-exist in the broadband access market. The network providers and venders are facing new challenges due to the emergence of a large quantity of bandwidth-intensive applications and services, in particular evolving from *Triple-play* (a bundle of services of voice, video, Internet) to *Quad-play* applications (integrated voice, video, Internet, and wireless services). Therefore, classical network equipments in some areas are aging and demand large-scale network reconstruction and/or upgrade.

Developing Trends of Broadband Access Networks

The (re)construction of broadband access networks reveals the following *trends* corresponding to the rapid growth of the future broadband access market.

- Fixed mobile convergence (FMC)

  FMC is expected to allow homogeneous access to heterogeneous network services without the constraints imposed by different technologies. In other words, an end-user can enjoy the same communication and Internet services anytime,

anywhere, and with any device. An emerging solution is to remove the barrier between various fixed and mobile technologies by the convergence in three areas, i.e., devices, services and networks [2–5].

- Simplifying the network structure

  Fast network (re)construction or upgrade is one of the keys to be successful in the broadband service market. Simpler network architecture can bring down the complexity of network planning and expedite the deployment with lower capital investment. In addition, it can decrease the difficulty in bandwidth control, Quality of Service (QoS) deployment, and reduce the operational cost for a network service provider.

- Incorporating new technical advances

  To achieve a cost-effective network deployment, employing state-of-the-art technical advances provides a practical and effective way to enhance system performance and QoS provisioning. On the other hand, new technologies lead to new network design problems when being adopted in practical scenarios. The techniques include advanced signal processing, tailoring system components (such as coding, modulation, and detection) for wireless environment, and different *diversity* techniques. Among these techniques, *diversity* is of primary importance in the wireless environment [6]. In particular, the recent advance in the spatial diversity technology of *wireless cooperative communication* has been recognized as one of the most effective paradigms to achieve higher system performance in terms of lower Bit Error Rate (BER), higher throughput, larger coverage, more efficient energy utilization, and higher network reliability [7, 8].

In this dissertation, we will focus on incorporating the wireless cooperative trans-

mission technology into the broadband access networks and exploiting the significant performance gains from the viewpoint of network planning.

1.2 Wireless Cooperative Transmissions

This research will explore the wireless cooperative transmissions in the following two contexts: single-cell cooperative relaying and multi-cell cooperative transmission (CT).

1.2.1 Single-cell Cooperative Relaying

With single-cell cooperative relaying, relay stations (RSs) are introduced into traditional cellular networks as shown in Figure 1.1, in which a long-distance low-rate/shadowed link can be replaced by multiple short-distance high-rate links. Thus the multi-hop relaying can extend the coverage of a cell and provide services for the Subscriber Stations (SSs) out of range of the cell. Also, the obstacles between a Base Station (BS) and SSs, which may impair the channel quality can be circumvented with the help of RSs. Most notably, due to the exploitation of all received signals that were originally taken as noise and interference, the wireless transmission rate and the system throughput/capacity can be substantially improved [7, 9]. From the perspective of reliability, multiple active wireless relay links can provide better robustness and fault tolerance than a single direct link from a BS to an SS. In a word, the data rate, the system capacity, the cell coverage, and the transmission reliability can be improved dramatically when cooperative relaying is employed in a cell.

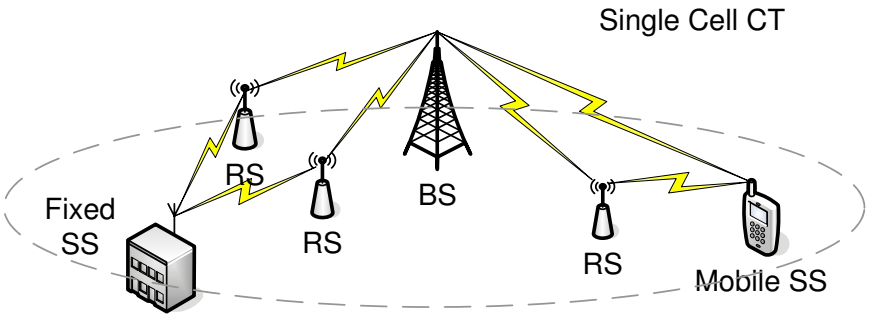


Figure 1.1: Single-cell cooperative relaying.

1.2.2 Multi-cell Cooperative Transmission

With multi-cell CT (also called inter-cell CT), a set of neighboring BSs collaboratively transmit data to an SS, as shown in Figure 1.2. The prerequisite of implementing multi-cell CT is a strong support of a central control unit, which could be an Optical Line Terminal (OLT) in a fiber optical network. The collaborative BSs can form a virtually distributed Multiple Input Single Output (MISO) array, which allows the implementation of downlink space-time coded transmissions to the SSs. By taking advantage of BS cooperative processing, multi-cell CT has exhibited a great potential on meeting the ever-increasing capacity demands for wireless communications. The BS cooperation can effectively combat inter-cell interference (ICI), which is the dominant impairment of a cellular network. Besides the ICI mitigation potential, other advantages of cooperative processing including the power gain, channel rank/conditioning advantage, and macro-diversity protection can also be addressed [8].

However, the performance gain of CT is highly dependent on the *locations* of RSs and/or BSs. In other words, if the RSs (BSs) are placed improperly, the receiver can not obtain a desirable transmission rate. The reason is that the radio signal attenuation between a transmitter and a receiver is a function of the propagation distance. Namely,

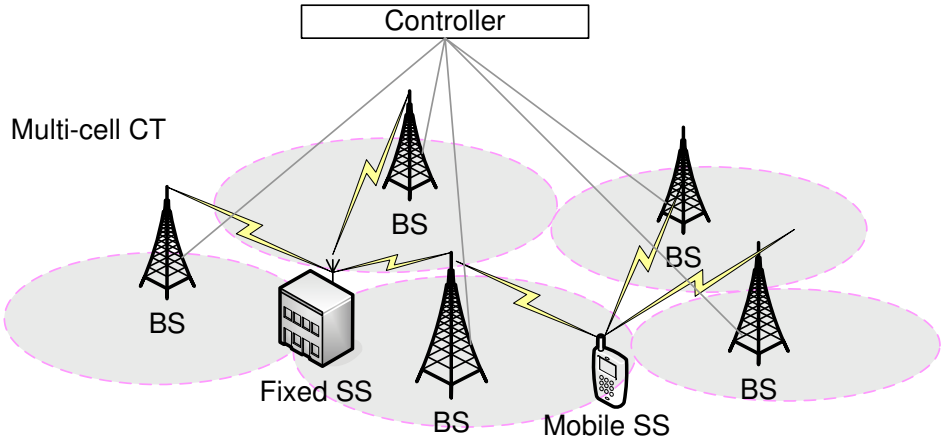


Figure 1.2: Multi-cell cooperative transmission.

the distance between the transmitter(s) and the receiver has critical influences to the receiver's achievable rate, and consequently, sum-rate or system capacity. And the system capacity often is a dominating factor to the operator's revenue. In the next section, we will further discuss the importance and challenges of location planning.

1.3 Importance and Challenges of Location Planning

Network design and planning has strategic implications for network operations. *Location planning problem* (or *placement problem*)[1] is the foremost critical issue for network planning and deployment, which has direct impacts on the subsequent QoS provisioning, especially the date rate and transmission delay.

Perhaps most important, a location decision could be irreversible in many situations since physical network structures are rather expensive to change sometimes [10]. The

[1]In this dissertation, the terms "location planning" and "placement" are interchangeable.

location planning imposes a fundamental influence on the network provider's long-term profitability and competitiveness. Poor placement may lead to a waste of capital expenditures (CAPEX) and unnecessary operation expense (OPEX), and consequently diminish the expected benefits.

Placement problems present rich applications in telecommunication domain for different aspects of location optimization, which is in general an NP-HARD problem [11]. In this research, the location optimization problem is also rather challenging due to a number of practical deployment considerations as follows.

- Many factors must be taken into account to achieve a good design, such as topology of the existing infrastructure, geographical distribution of traffic load demands, and terrain topography, etc. Whether a particular candidate position is suitable for placement depends largely on which factors are considered, thus we must thoroughly analysis the impact of influencing factors on the placement.

- The determination of optimal placement is closely related to the network operation with respect to the specific technologies employed in the broadband access networks, especially the cooperative transmission based physical layer technologies and medium access control (MAC) layer schemes. Thus, it is a challenging task to capture the characteristics of the network model precisely with a mathematical formulation for the placement problem.

- The placement problem and the resource allocation problem are usually coupled. In other words, the optimality of system performance is achieved when both locations and resource allocation are jointly determined optimally. It will no doubt increase the difficulty of problem formulation and the development of solutions.

1.4 Research Motivations, Objectives and Contributions

Research Motivations

In practice, with the modern network's evolution and growth, network planning is a never-ending process. A new round of work may be triggered due to the following deployment demands.

1. The needs of boosting wireless network throughput/capacity and/or extending mobile service coverage. For instance, low signal-to-noise-ratio (SNR) at cell edge is expected to be improved to achieve the ubiquitous mobile broadband access; subscribers in densely populated area may have higher traffic demand and QoS requirements.

2. The change of wireless network environments. Typically, new shopping centers and communities may lead to new hotspots or new (micro)cell sites; newly constructed buildings may change the multipath environments of the existing telecommunication systems, possibly resulting in "black holes" in the coverage area due to the shadowing effect.

3. Stringent traffic demands. For instance, spontaneous events such as big conferences, sports meetings, and fairs may launch significant traffic demands during the period of the events.

4. The need of constructing a new metropolitan-area broadband access network for residential, business and mobile access services.

Solving the above problems will quickly promote potential subscriber growth and enhance the competitiveness of network service providers.

Research Objectives and Contributions

The objective of this dissertation is to tackle the task of strategic location planning in practical deployment scenarios so as to meet the requirements of broadband access services. The practical deployment environment can be classified into two categories: one is an existing wireless access network; the other is a prospective broadband access network. Inspired by the significant performance improvements due to CT, we will incorporate this promising technical advance into the network planning to achieve a more cost-effective design.

For the first category of deployment environment, to address the problems in a cell (e.g., the aforementioned demands 1, 2 and 3), instead of setting up a new cell by deploying a BS like the way in traditional cellular networks, we introduce RSs to achieve the same design criteria. The RSs can be designed and equipped with either simple or complex hardware and software with respect to the type of adopted antennae and intelligent control functions (e.g., admission control, congestion control, adaptive modulation and coding, authentication and authorization, scheduling, joint RS/MS roaming and routing), which are decided according to the network topology, traffic demands, user mobility pattern, and QoS requirements [12–15]. Moreover, RSs are often with lower costs, and can be deployed more rapidly and easily when compared with BSs. Thus, RSs have been recognized as a competitive solution for network providers due to the economical advantages and the appealing performance characteristics through cooperative relaying. We study the problem of optimal RS placement in wireless cooperative networks. To the best of our knowledge, this is the first study on RS location planning,

which incorporates the CT technology to achieve better performance. Specifically, the contributions of the research in the first deployment environment are summarized as follows.

- An RS placement problem is studied with the objective of cell capacity maximization given a specific number of RSs to be deployed. Relay allocation and bandwidth allocation are jointly optimized.
- A dimensioning and location planning problem is investigated with the objective of minimizing the total cost of deploying RSs within a cell, in which multi-level cooperative relaying is employed. Besides RS placement, relay allocation and signal relay sequence design are jointly optimized.
- The optimization frameworks are provided to capture the characteristics of the above two problems with mathematical formulations, respectively.
- The solution algorithms are developed to obtain the optimal (or near-optimal) results in polynomial time.
- Numerical analysis and case studies are conducted to validate the performance benefits due to RS placement.

For the second category of deployment environment, to realize the ambition of FMC [2], incorporating state-of-the-art optical and wireless networks has been envisioned as a norm. Among all the wire-line networks, Passive Optical Networks (PONs) have attracted extensive interest from both academia and industry, such as Gigabit PONs (GPONs) [16] and Ethernet PONs (EPONs) [17, 18]. Due to the nature of all-passiveness, namely, no active electronics are required within the access networks, PONs provide the most cost-effective solution for wire-line networks in terms of both

CAPEX and OPEX, without sacrificing transmission bandwidth. Although PONs have advantages over Active Optical Networks (AONs), the CAPEX and OPEX of PONs are still much more than those of a wireless access network. Moreover, similar to all the other wired-line networks, a PON lacks mobility support.

It has been taken as a promising approach to integrate a PON with a wireless access network (e.g., an IEEE 802.16-based WiMAX network [19, 20], an LTE cellular network). The integration aims at exploiting their major strengths and complementing each other. It is notable that an optical network can yield high bandwidth, high reliability, but at the expense of costly deployment. On the other hand, wireless networks demonstrate flexible and cheap deployment, excellent mobility support, but subject to the limitation of scarce and expensive frequency spectrum. In the integrated architecture, the wireless network serves as a front-end access point and the PON is desirable to be the backhaul.

Unlike the traditional hybrid network architecture, in which the wired and wireless networks are just interconnected through a common standard interface (e.g., Ethernet) [21, 22], an integrated PON-wireless cooperative network (WCN) architecture is investigated. A key technology of multi-cell CT is incorporated. We study the problem of dimensioning and location planning for an integrated PON-WCN by exploiting both the performance benefits due to CT and the interplay advantages of wired/wireless network integration. Specifically, the research contributions in the second deployment environment are summarized as follows.

- An optimization framework in the integrated PON-WCN is provided, which considers placement, BS-User association, and resource breakdown assignment together in a single stage.

- A solution approach is proposed for the complex optimization problem so that it

is reformulated from an unsolvable Mixed Integer Nonlinear Program (MINLP) to a solvable Mixed Integer Linear Program (MILP).

- Numerical analysis and case studies are conducted on the performance and economic benefits of the proposed framework.

1.5 Outline of the Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 presents some background knowledge and literature survey, where the issues of wireless cooperative transmission, location planning in telecommunication networks, and integration of optical and wireless networks are reviewed. The optimal RS placement is studied in Chapters 3, in which the design objective is cell capacity maximization. Chapter 4 extends the work of RS placement to multi-level cooperative relaying. Both Chapters 3 and 4 address the location planning in an existing wireless access network under single-cell CT. Chapter 5 investigates the dimensioning and site planning problem in an integrated PON-WCN architecture under multi-cell CT. For each placement problem in Chapters 3 to 5, the corresponding mathematical formulations, solutions and numerical investigation are provided, respectively. Finally, conclusions and future work for this research are given in Chapter 6.

Chapter 2

Background and Literature Review

This chapter reviews important references from a broad array of literature that relate to the problems studied by the dissertation. Our main objectives are to make the readers aware of the many considerations involved and the previous related research, and also to facilitate the development and discussion of the following chapters in this dissertation.

2.1 Cooperative Transmissions in Wireless Networks

The broadcast nature of wireless communications suggests that a source signal transmitted towards the destination can be overheard at neighboring nodes. Instead of treating this overheard information at the surrounding nodes as "interference", cooperative communication allows to exploit this overheard information and retransmit towards the destination to create spatial diversity, thereby to obtain higher throughput and reliability. In the following, some background and related work are presented from the perspective of single-cell cooperative relay transmission and multi-cell cooperative transmission, respectively.

- **Single-cell cooperative relay transmission**

From information theory research, the early study for wireless relay channel capacity was explored in the work by Cover and El Gamal [23], which focused on the simplest relay channels consisting of a source, a destination, and a relay. As shown in Figure 2.1, based on the assumption that all nodes operate in the same waveband, the system can be taken as a hybrid broadcast channel (BC) from the viewpoint of the source while a multiple access channel (MAC) from the viewpoint of the destination.

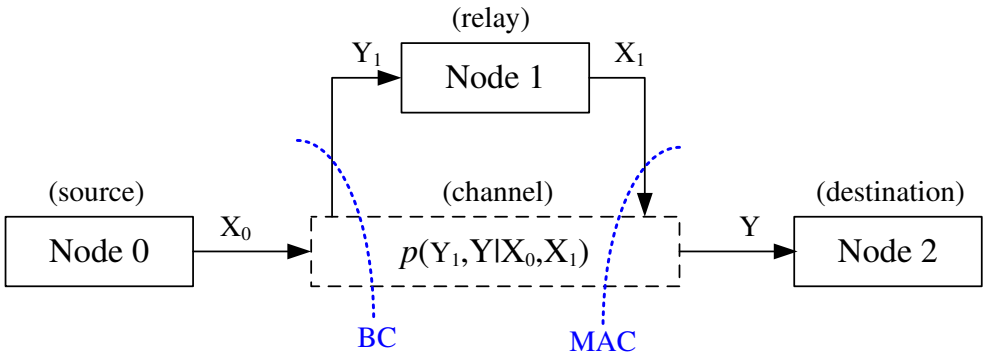


Figure 2.1: The BC/MAC in three-node relay channel model

Cover and El Gamal derived the following maximum achievable rate for a relay channel [23]:

$$R < \max \min\{I(X_0; Y_1, Y|X_1), I(X_0, X_1; Y)\} \tag{2.1}$$

where the first term in $\min(\cdot,\cdot)$ represents the overall rate from the source node to the relay node and the destination node, which corresponds to the BC part; while the second term represents the overall rate from the source node and relay node to the destination node, which corresponds to the MAC part. Eq. (2.1) exercises max-flow min-cut in the calculation, where the rate of the information flow transmitted on the

relay channel is constrained by the bottleneck corresponding to either the first cut (BC) or the second cut (MAC).

Based on the approaches of relay processing and re-transmission, there appear to be three general classes as follows.

1. Amplify-and-forward (A-F): The relay acts as analog repeaters, and transmit the amplified version of its received signal. Although noise is also amplified by cooperation, the destination receives multiple independently faded versions of the signal and can make better decisions on the detection of information.

2. Decode-and-forward (D-F): The relay demodulates and decodes the source's data packets, then forwards it to the destination using a different code. The relay cooperates with the source in the transmission. It is the only one where interference completely disappears. The D-F scheme realizes the *transmitters cooperation*.

3. Compress-and-forward (C-F): The relay compress the estimated symbols of the signal received from the source and transmit a compressed version to the destination. The C-F scheme is especially suitable for the situation where the channel between source and relay is worse than that between source and destination.

Multiple relay extensions have also been examined [7, 9, 24, 25]. Xie *et al.* revealed the "scaling law" of the transport capacity and the optimal operation in a wireless network [9]. In [7], an achievable rate formula and a simple coding scheme to achieve it were presented for the multiple level relay channel. A multi-source multi-relay coding scheme was developed in [24]. Gastpar *et al.* developed two coding strategies for a multiple-relay channel which mimicked multiple-antenna transmission and multiple-antenna reception, respectively. The strategies can achieve capacity when the terminals

formed two closely-spaced clusters [25].

For cooperative relay communications, the resource allocation is concerned with more than one hop communication, such as power allocation [26] and rate allocation [27]. Himsoon *et al.* [26] studies the issue of lifetime maximization by cooperative-node employment and relay deployment in wireless networks. The optimization problem focused on determining which nodes should cooperate and how much power should be allocated for cooperation. Kim *et al.* [27] proposed an optimal resource allocation strategy for cooperative relaying-enabled OFDMA multi-hop wireless networks. A cross-layer optimization problem was formulated that maximizes the balanced end-to-end throughput under the routing and the PHY/MAC constraints. In addition, the research on optimizing relay assignment have become another focus, since relay nodes can also be considered as system "resource" which should be optimally allocated so as to achieve certain QoS requirements for subscribers (e.g., [28–32]).

With cooperative relay, the source node and the relay node(s) virtually form an antenna array which provides the potential spatial diversity. In other words, the geographically separated transmitters (i.e., the source and the relay nodes) independently transmit and/or process (semi) copies of information sent by the source, while the receiver is capable of combining the received information using advanced signal processing techniques to create an estimate of the transmitted information. Thus cooperative diversity can effectively combat the sever signal attenuation due to fading [33].

Because of these large benefits of cooperative relay techniques, several standardization groups, such as IEEE 802.16 and IEEE 802.11, have started standardization processes to include such technologies into their prevailing standards. For instance, IEEE 802.16j [12], the Multihop Relay Specification for 802.16, is under development by the IEEE 802.16's Mobile Multihop Relay (MMR) Study Group currently.

- **Multi-cell cooperative transmission**

There has been increasing interest to study the cooperative processing between BSs with a cellular system [8, 34–38], which is also referred to as multi-cell cooperative transmission in this dissertation. Zhang *et al.* in [8] revealed a great potential of BS cooperative processing on capacity enhancement based on information-theoretic dirty paper coding (DPC) approach and practical signal-processing standpoints. A multi-cell cooperative zero-forcing beamforming (ZFBF) scheme was proposed in [34], which combined with a simple sub-optimal users selection procedure for the Wyner downlink channel setup. The per-cell sum rate of ZFBF was proved to experience the same growth rate as the optimal DPC scheme asymptotically. Shao *et al.* proposed a multiple-input-multiple-output (MIMO) scheme that can effectively combat co-channel interference with only local BS coordination and retain the high peak rate achievable for point-to-point single-cell communications [35]. Garcia *et al.* investigated how coverage, handover and intercell-interference evaluation metrics were redefined for CT, and the results showed that CT permitted superior network designs with higher traffic, greater overlap of cell coverage areas and required less BSs to meet the traffic demand [36]. Tolli *et al.* studied the downlink resource allocation problem with differen BS power constraints with the objective of weighted sum rate maximization in a cooperative MIMO-OFDM cellular system. The soft handover (SHO) region was defined for user with similar received power levels from adjacent distributed BS antenna. Results demonstrated the significant performance gains in terms of both the transmission rates for the users within the SHO region and overall system capacity (weighted sum rates) [37]. The theoretical system capacity of a cooperative multi-cell downlink transmission was evaluated using numerical optimization for different BS and MS antenna configurations [38].

2.2 Location Planning in Communication Networks

The node placement problem in communication networks is rooted in the location problems in operations research. The location problems study the quantifiable relationship between facility locations and the performance of a system in which the facilities are located [39]. Location problems and models may be classified mainly in two ways as follows.

First, they may be classified based on the topography (e.g., continuous versus discrete location problems). In continuous location models, both demands and facilities may be located anywhere on a plane. The demands can be represented by a spatially distributed probability distribution. In many cases, these problems can be cast as convex optimization problems and approaches such as Lloyd-Max algorithm [40] were employed to solve them. The discrete location models allow a broader range of problems in which the demands and facilities are assumed to occur only in sets of nodes and links. Since the structure of underlying network is lost, these problems are even more difficult to solve, especially for large-scale problems.

Second, the location problems may be classified according to the number of facilities to be located. Some location problems has the constraint that the number of facilities is exogenously specified, such as *p-median* [41], *p-center* [42], and *maximum covering* [43] problems. In other cases, the number of facilities is a model output (objective), such as *set covering* [44] and *facility location* problems [45].

Generally, we can classify the communication networks as wire-line networks and wireless (mobile) networks. In the following, we provide an overview of recent work addressing location problems in different types of networks.

- **Location planning in wired networks**

Regarding wired networks, there are optical networks using fiber and 'classical' telephone/cable networks using coaxial cables.

With respect to optical networks, the use of wavelength-division multiplexing (WDM) led to the *converter* placement problem. WDM is a technology which multiplexes multiple optical carrier signals on a single optical fiber by using different wavelengths of laser light to carry different signals. Thus when and where in the network should we optimally switch from one wavelength to another in order to achieve a certain criterion? Ruan *et al.* investigated the problem of finding minimum set of network nodes to place converters to support broadcast in the WDM optical networks [46], in which two approximation algorithms were given by studying two related problems, namely, color-covering and vertex color-covering problem. Jia *et al.* studied the problem of placing the minimal number of converters in a multihop WDM network, and the objective was to achieve equal number of wavelengths and maximal link load [47]. Chu *et al.* considered wavelength converter placement and routing and wavelength assignment together and aimed at minimizing the overall blocking probability [48].

The need to maintain a consistent quality of a signal over long fiber links led to the problem of *amplifier* placement. In [49], Tran *et al.* presented two amplifier placement methods to minimize the number of amplifiers in metropolitan WDM rings based on integer programming techniques. The amplifiers were required to operate the self-healing ring under a normal or any single-link or single-node failure conditions. In [50], Rocha *et al.* demonstrated a simple way to improve the OMEGA (Optical Metro network for Emerging Gigabit Applications) testbed performance, which is to place amplifiers at the node input ports. The authors measured a consistent improvement of 2dB in power budget. The gained margin could hence be used for adding functionality to a network

with standard fibre spans that are representative of real optical metropolitan networks.

The situation gets even more complicated when multicast applications emerged. The problem, then, is to decide on placement of *splitters* that allow simultaneous transmissions to all the destination nodes. Lin *et al.* studied the splitter placement problem in all-optical WDM networks in which a light forest was used to realize a multicast connection. The design objective was to minimize the average per link wavelength resource usage of multicast connections given the number of splitters [51]. In [52], Hu *et al.* addressed the problem of reducing the number of splitters on a directed fiber tree without serious performance degradation. Routing and wavelength assignment algorithm was also proposed to maximize requests on a directed fiber tree.

- **Location planning in wireless networks**

In wireless access networks, placing network equipments (e.g., BSs, access points, wireless gateways, wireless sensor nodes, relay nodes) also plays a key role in the overall network performance.

In cellular networks, coverage planning is often the main objective for BS placement. In [53], Amaldi *et al.* evaluated the number, location, and transmission powers of BSs in a 2-D urban setup. In [54], Weicker *et al.* followed a heuristic approach based on an evolutionary paradigm to solve the problem of BS placement with the objective of minimizing both the cost and the channel interference. And a major finding of [54] was a strong influence of the choice of the multiobjective selection method on the utility of the problem-specific recombination leading to a significant difference in the solution quality.

In traditional wireless sensor networks (WSNs), there are two types of nodes, i.e., sinks (or BS), and sensor nodes (SNs). A large number of these sensors can be net-

worked to perform various applications, such as failure diagnosis, habitat, environment monitoring or military applications. The lifetime of the network that is determined by the power constraint of each sensor. The energy issue has always been one of the main concerns in the WSNs design. Recently, relay nodes (RNs) have been introduced in WSNs which are also referred to as cluster-based WSNs. In this case, the RNs perform data aggregation and fusion so as to decrease the number of hops to the sink. The objective of placing RNs is to provide a better balance of the energy depletion and prolonged lifetime of the SNs [55] [56].

In WSNs, the coverage of each sensor is still a key consideration to ensure network connectivity. Some studies formulated the node placement problem by making an analogy to the *illumination* or *art gallery* problems [57]. The *illumination* problem is to study how many light sources are needed so as to illuminate the interior of a given polygon, and the *art gallery* problem rephrases the same problem by asking how many security cameras are needed in order to monitor the interior of an art gallery with a given floor plan. In [58], SNs are assumed to be energy constrained while the RNs are not. The task of RN placement is modeled as a *minimum set cover problem* that can be optimally solved by a recursive algorithm. In [59], both SNs and RNs are energy limited, where the RNs function as Cluster Heads (CHs) each governing a number of SNs. Based on the Far-Near Max-Min principle, an enhanced localized heuristic scheme was proposed to yield an optimal or near optimal solution. In [60], a number of deployment strategies for determining optimal placement of SNs, RNs, and sinks, were introduced, in order to achieve the required coverage, connectivity, bandwidth, and robustness. The placement problems for both reliable and unreliable/probabilistic detection models were formulated as Integer Linear Programs (ILPs). In [56], based on data aggregation through competitive machine learning, a cluster-based self-organizing

data aggregation algorithm was proposed.

In WLANs, extending limited coverage and increasing network capacity are usually the main objectives. The study in [61] introduced an extension point (EP) placement algorithm which aimed at improving the throughput of a rectilinear network by way of a divide-and-conquer searching process. Similarly, to achieve optimal placement for a fixed number of RNs for relaying traffic in a two-hop fashion in WLANs, the problem was formulated as a *p-median* problem in [62] [63]. A solution based on Lagrange relaxation and sub-gradient optimization algorithms was presented.

In wireless mesh networks (WMNs), Internet Transit Access Points (ITAPs) were employed in [64] to serve as gateways to the Internet. These ITAPs formed a wireless neighborhood network of the wired Internet. Algorithms were proposed to make informed placement decisions based on the neighborhood layouts, where the fault tolerance and workload variation were taken into account. In [65], So *et al.* formulated the relay placement problem in heterogeneous WMNs as an integer programming (IP) which was solved by Bender's Decomposition algorithm. In [66] and [67], for coverage/topology planning in WMNs, Amaldi *et. al* presented a novel optimization model and a relaxation-based heuristic for large-size network instances. The issues of traffic routing, interference, rate adaptation, and channel assignment were also considered.

To solve the location optimization problem, numerous techniques have been used, e.g., iterative methods (e.g., quasi-Newton in [68], least square and linear regression in [69]), pruning-searching techniques (e.g., Hooke-Jeeves in [68], Nelder-Mead in [70]), combinatorial optimizers (e.g., genetic algorithm in [71] [15] [54], tabu search in [72]), and heuristics algorithms [62, 65–67]. Each solution is problem-specific and needs careful design based on the unique characteristics of the location problem.

2.3 Integration of Optical and Wireless Networks

Wireless and optical integration is pivotal to the exploitation of future access solutions. Optics provides greater bandwidths and long reach while wireless offers easier deployment, mobility and freedom to end users. By combining both, service experience is highly enhanced. Previous research have been reported on integrating optical with wireless technologies, including optical millimeter-wave (mm-wave), Ultra-Wide Band (UWB), WiMax, WLAN, etc.

In [73], Jia *et al.* discussed several key enabling technologies for hybrid optical-wireless access networks, including mmWave generation, upconversion, and transmission in a downlink direction, and full-duplex operation based on wavelength reuse by using a centralized light source in an uplink direction. In [74], Bock *et al.* presented an architecture in which a WDM/PON fibre link run to each home, and then was split into several UWB access points to give wireless connectivity inside the home. In [19], a framework on integration of PON and WiMAX for broadband fixed/mobile convergence access was proposed, in which the research issues on the media access control (MAC) layer and integrated control plane were elaborated. In [75], Chaudhry *et al.* developed an application-controlled handover scheme, which can keep channel continuity in the wired-wireless synergy network environment that consists of 3G Universal Mobile Telecommunications Systems (UMTS), WLAN, Wireless Personal Area Networks (WPAN) (e.g., UWB) and optical fibre network. The simulation results showed an enormous achievement, including reduced the packet drop, Bit Error Rate (BER), response time and power consumption.

Other investigations on the fiber-wireless (FiWi) access networks were conducted in [21] [22], where a BS can send data to gateways/ONUs over intermediate wireless BSs instead of directly over an ONU. These research focused on the placement of wireless

BSs or access points, routing algorithms, and performance evaluation [21] [22].

For the synergy of wireless and optical fiber communications, the radio-over-fiber (RoF) technology plays a significant role. RoF refers to a technology whereby light is modulated by a radio signal and transmitted over an optical fiber link to facilitate wireless access [76]. The use of RoF was first proposed and demonstrated in 1990 by Cooper [77]. Since then RoF technology has gained great monument from international research community. With the extensive use of microcells in cellular systems, it is needed to interconnect huge number of (micro)cells using optical fibers, which offers a high transmission capacity at low cost. RoF technology has further benefits in terms of flexible network channel allocation and rapid response to traffic demand variations since frequency allocation can be executed at a central station. Lin *et al.* proposed a dynamic wavelength-allocation scheme at the bursty traffic load for WDM fiber-radio ring access networks in [78]. Pfrommer *et al.* [79] demonstrated simultaneous wireline (600 MHz) and wireless (5.5 GHz) data transmission in a hybrid fiber-radio access network over cable-service interface specification. A scheme for quantizing radio signals over fiber was investigated in [80]. Liu *et al.* provided an overview of cost-effective wireless-over-fiber technology in [81].

2.4 Summary

This dissertation is built significantly from the fundamental work on CT. Thus this chapter has began with some background and related research on CT. The studies on relay channel, the extension of multi-relay networks, and the recent research on multi-BS cooperative processing have been reviewed. *Section 2.2* has provided an overview of the research on location problems. The various applications and design objectives of location problems have exhibited great differences. Thus, each placement problem

requires careful and thorough analysis and modeling to capture the specific characteristics. Finally, *Section 2.3* has reviewed recent work on optical-wireless integration, which as aforementioned is a developing trend for future BAN. *Chapter 5* is inspired by previous research on the optical-wireless integration and the advances in multi-cell CT, aiming to achieve a more cost-effective deployment for FMC.

New technologies will lead to *new* models and, consequently, to *new* location problems and solutions. To the best of our knowledge, few research have incorporate the advances in CT into location planning. In the following chapters, we will study the placement problems in practical deployment scenarios, specifically, in an existing wireless access network (*Chapters 3* and *4*) and a perspective integrated optical-wireless network (*Chapter 5*), respectively.

Chapter 3

Capacity Enhancement with Relay Station Placement

As aforementioned, cooperative relaying technology provides a competitive solution to capacity/throughput enhancement. From the viewpoint of network service provider, achieving maximum capacity (revenue) with a certain CAPEX on network equipments (e.g., relay stations) is always the goal of deployment. This chapter studies the **C**apacity **M**aximization problem by way of **R**S **P**lacement (also called **CMRP** problem) in an existing broadband wireless access network. We will concentrate on location planning in a single cell in *Chapters 3* and *4*. And the solution can be easily extended to multi-cell scenarios. The cell capacity is defined as the total achievable rate at each subscriber station (SS). Bandwidth allocation is taken as another closely related design dimension, which is jointly optimized for better performance.

Note that in this research two types of RSs, namely fixed RSs and nomadic RSs, could be used for delployment. Since RSs can not be placed anywhere, certain outdoor physical locations are taken as candidate positions (CPs) for placing RSs permanently (for fixed RSs) or for a period of time (for nomadic RSs). These CPs are required

to provide uninterrupted power supply. Nomadic RSs are mainly used to mitigate capacity/coverage problems and improve adaptability to traffic demands growth due to special events, emergency response or newly emerging hot spots that generate large traffic demands. Thus, an efficient RS placement computation method which allows timely adaptation to traffic demand change is needed.

3.1 System Model

3.1.1 Relay-Based Wireless Network Architecture

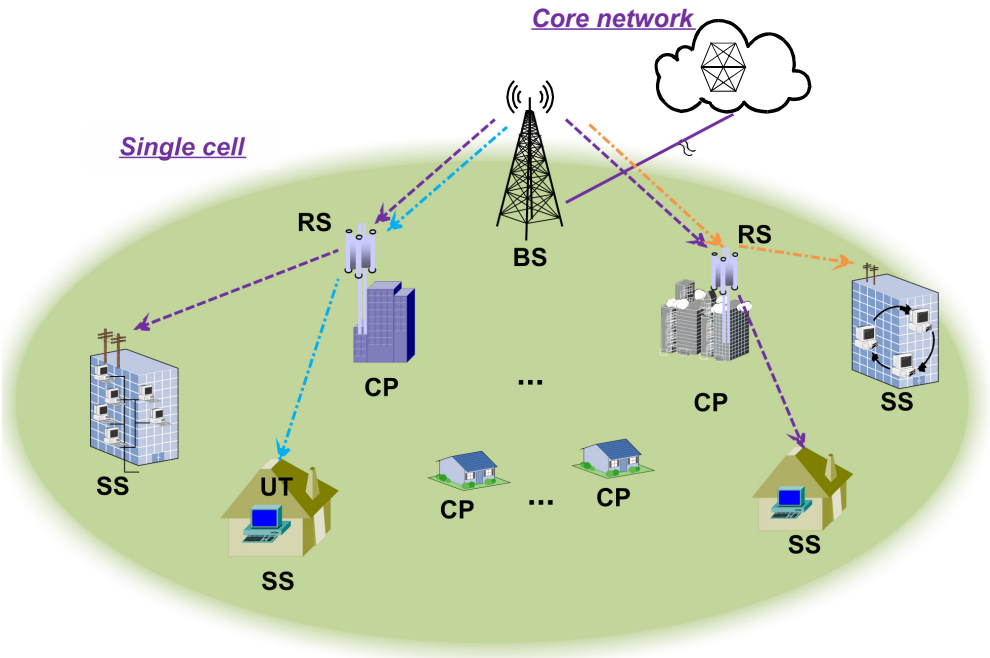


Figure 3.1: A relay-based broadband wireless access network architecture.

In this study, a three-tier network architecture is considered as shown in Figure 3.1, which consists of the following network entities: the BS, the RSs, and the fixed SSs. The BS serves as a central controller/coordinator and handles all the routing and

signaling issues in the cell. The RSs are responsible for relaying data between the BS and the associated SSs based on a given cooperative relaying strategy. The RSs have no direct connections to the core network and are eligible to be deployed at certain outdoor candidate positions (CPs). The SS may stand for a hotspot or a building at which a large amount of traffic load generated by user terminals (UTs) are aggregated at the corresponding SS. To a certain degree the network design and deployment depends on the geographical distribution of traffic demands. Based on the statistical data analysis of traffic measurement and monitoring [82], the mean and peak traffic load demand of each SS can be estimated for the network planning. For the sake of long-term network deployment, upgrading and extensions, we take the geographical distribution of peak traffic load demands plus the anticipation of traffic load growth at SSs as a known input in this research.

The BS can be multiple accessed simultaneously by different SSs at their assigned frequency band with the Orthogonal Frequency-Division Multiple Access (OFDMA) technique. In other words, each transmission between the BS and an SS is inherently an instance of the basic "BS-RS-SS" three-node relay model, where the three wireless links, BS-SS, BS-RS and RS-SS links, share a common frequency spectrum. The COST231-Hata model is adopted as the radio propagation model which is applicable to the transmissions inside an urban environment [83]. Small scale fading is not explicitly included in the system model since a long-term planning and design is targeted.

In addition, only two-hop cooperative relaying is assumed in this chapter. The more complex issue of RS placement with multi-hop relaying is studied in *Chapter 4*.

3.1.2 Cooperative Relaying Strategy

A cooperative relay with the D-F technology is adopted in the wireless networks. The relay cooperates with the source by demodulating and decoding the received data packets, and forwarding them to the destination possibly using a different code. With D-F cooperative relaying, an alternative connection between the source and the destination is established through two shorter-ranged links with better reliability and larger data rates. D-F allows an interference-free transmission when all transmissions share a common frequency band [7]. The simplest relay channel model is a class of three node communication networks [23], as shown in Figure 3.2. The achievable rate for the destination node is:

$$r = \min\{C(\frac{\theta P_s d_{sr}^{-\alpha}}{N_r}), C(\frac{P_s d_{sd}^{-\alpha} + P_r d_{rd}^{-\alpha} + 2\sqrt{(1-\theta)P_s d_{sd}^{-\alpha} P_r d_{rd}^{-\alpha}}}{N_d})\} \tag{3.1}$$

where α is the path loss exponent, P_s and P_r are the transmit power of the source and the relay, respectively. d_{sr}, d_{sd} and d_{rd} are the distances illustrated in Figure 3.2. N_r and N_d are the noise power at the relay and destination node, respectively. $C(.)$ is a Shannon function defined as $C(x) = \frac{1}{2}\log(1+x)$ for $x \geq 0$. Parameter θ $(0 < \theta < 1)$ is the transmit power allocation ratio at the source node between the "source-relay" path and "source-destination" path. θ is an important parameter that can affect the achievable rate dramatically.

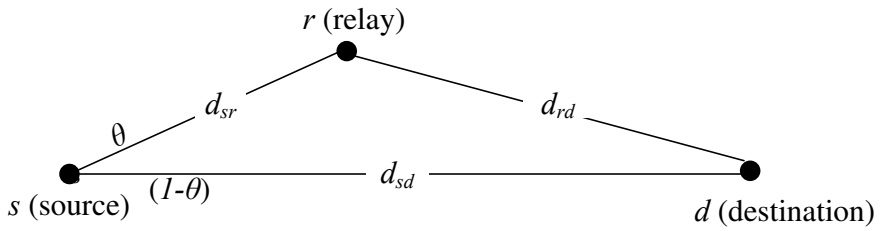


Figure 3.2: The three-node relay channel model.

To facilitate the CMRP formulation in a wireless network with multiple RSs and SSs, we need to derive a close form expression of the optimal power allocation ratio in the fundamental three-node relay model. Let θ^* denote the optimal power allocation ratio at the source node. For simplicity, we assume $N_r = N_d = N_0$. Given the locations of the source, relay and destination, to maximize the achievable rate r, θ^* is expressed as

$$\theta^* = \arg\max_{\theta}[\min(r^{(1)}(\theta), r^{(2)}(\theta))] \tag{3.2}$$

where $r^{(1)}(\theta) = C(\theta\frac{P_s}{N_0 d_{sr}^{\alpha}})$ is a monotonously increasing function of θ, and $r^{(2)}(\theta) = C(\frac{P_s}{N_0 d_{sd}^{\alpha}} + \frac{P_r}{N_0 d_{rd}^{\alpha}} + \frac{2}{N_0}\sqrt{(1-\theta)\frac{P_s P_r}{(d_{sd} d_{rd})^{\alpha}}})$ is a monotonously decreasing function of θ. The optimal value of θ must be the value when $r^{(1)}(\theta) = r^{(2)}(\theta)$. Therefore, θ^* is given by

$$\theta^* = \begin{cases} \frac{d_{sr}^{\alpha}}{P_s}(\frac{2(-d_{rd}^{\alpha}P_r + \sqrt{u})}{d_{sd}^{\alpha} d_{rd}^{\alpha}} + \frac{P_s}{d_{sd}^{\alpha}} + \frac{P_r}{d_{rd}^{\alpha}}) & \text{for} \quad u \geq 0 \\ 1 & \text{for} \quad u < 0, \end{cases} \tag{3.3}$$

where $u = P_r(d_{sd}^{\alpha} - d_{sr}^{\alpha})(P_s d_{rd}^{\alpha} - P_r d_{sr}^{\alpha})$.

3.2 Problem Formulation

3.2.1 An Mixed Integer Nonlinear Model

The RS placement problem with multiple RSs and SSs is studied in this section, where the full-duplex D-F cooperative relaying strategy is adopted. Note that the other cooperative relaying strategy (e.g., the C-F scheme) can be adopted, and the RS placement problem can be formulated in a similar manner.

Definition 3.1. Problem CMRP. Given a set of fixed number of RSs, the geographical distribution of the traffic load demands of SSs, the locations of CPs for

deploying RSs, the total bandwidth allocated to the cell, and the transmission power of the BS and RSs, the goals of the Capacity Maximization RS Placement (CMRP) problem are to

(1) obtain the maximum cell capacity, denoted by $\mathcal{C}$,

(2) determine the optimal locations of these RSs,

(3) assign serving RS(s) to each SS,

(4) and allocate bandwidth to each SS.

We define the decision variables of Problem CMRP as follows. $\mathbb{Z} = (z_m)_{1\times M}$ is an *RS location incidence vector*, where

$$z_m = \begin{cases} 1, & \text{if an RS is placed at CP}_m, m \in \mathbb{N}_{CP}; \\ 0, & \text{otherwise.} \end{cases}$$

$\mathbb{A} = (a_{mn})_{M\times N}$ is an SS-CP incidence matrix, or the *location-allocation matrix*, where

$$a_{mn} = \begin{cases} 1, & \text{if SS}_n \in \mathbb{N}_{SS} \text{ is relayed via an RS located at CP}_m \in \mathbb{N}_{CP}; \\ 0, & \text{otherwise.} \end{cases}$$

$\mathbb{W} = (\omega_n)_{1\times N}$ is the *bandwidth allocation vector*, where ω_n is the amount of bandwidth assigned to SS_n. The notations taken in Problem CMRP are listed in Table 3.1.

The formulation of Problem CMRP can be expressed as:

$$\text{(CMRP)} \underset{\mathbb{Z},\mathbb{A},\mathbb{W}}{maximize} \quad \mathcal{C} = \sum_{n=1}^{N} \omega_n \sum_{m=1}^{M} a_{mn} \min[C(\theta^* \frac{P_{BS}}{d_m^\alpha}), C(\frac{P_{BS}}{d_n^\alpha} + \frac{P_{RS}}{d_{mn}^\alpha} + 2\sqrt{\overline{\theta^*} \frac{P_{BS}P_{RS}}{(d_n d_{mn})^\alpha}})] \tag{3.4}$$

subject to:

$$\omega_n \sum_{m=1}^{M} a_{mn} \min[C(\theta^* \frac{P_{BS}}{d_m^\alpha}), C(\frac{P_{BS}}{d_n^\alpha} + \frac{P_{RS}}{d_{mn}^\alpha} + 2\sqrt{\overline{\theta^*} \frac{P_{BS} P_{RS}}{(d_n d_{mn})^\alpha}})] \geq \rho_n \qquad \text{for} \quad \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.5}$$

$$\sum_{m=1}^{M} a_{mn} = 1 \qquad \text{for} \quad \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.6}$$

$$a_{mn} \leq z_m \qquad \text{for} \quad \forall n \in \mathbb{N}_{\mathbb{SS}}, \forall m \in \mathbb{N}_{\mathbb{CP}} \tag{3.7}$$

$$\sum_{m=1}^{M} z_m = K \tag{3.8}$$

$$\sum_{n=1}^{N} w_n \leq BW, \tag{3.9}$$

$$a_{mn} \in \{0, 1\} \qquad \text{for} \quad \forall n \in \mathbb{N}_{\mathbb{SS}}, \forall m \in \mathbb{N}_{\mathbb{CP}} \tag{3.10}$$

$$z_m \in \{0, 1\} \qquad \text{for} \quad \forall m \in \mathbb{N}_{\mathbb{CP}} \tag{3.11}$$

where d_m (d_n) represents the distance between the BS and CP_m (SS_n), and θ^* is the optimal power allocation ratio which is given by (3.3).

The objective function (3.4) maximizes the cell capacity. Constraint (3.5) ensures the throughput of each SS is no less than its minimal traffic load demand. Constraint (3.6) ensures that each SS is exclusively associated with a single RS. Constraint (3.7) ensures that if SS_n is associated with CP_m, an RS must first be placed at CP_m. Constraint (3.8) stipulates that K RSs are allocated among the M CPs. Constraint (3.9) is the bandwidth constraint of the cell. Constraint (3.10) and (3.11) state that each entry in the decision variables of $\mathbb{A}$ and $\mathbb{Z}$ is binary.

The mixed-integer and nonlinear characteristics of the formulation will result in computational intractability. Then we will reformulate Problem CMRP from an MINLP into an MILP such that the optimal solution can be obtained by using a well recognized MILP solver CPLEX [84]. To facilitate a more systematic and efficient computation,

we will propose heuristic solutions, which will be elaborated in Section 3.3. The optimal result obtained by CPLEX will be taken as a benchmark to evaluate the proposed heuristic solutions.

3.2.2 An Mixed Integer Linear Model

The reformulated problem is denoted as Problem CMRP-MILP. In order to realize the model transformation and eliminate the nonlinearity in the original CMRP model, a new set of variables is introduced and defined as follows: $\mathbb{B} = (b_{mn})_{M\times N}$ such that

$$b_{mn} = \begin{cases} \omega_n, & \text{if } SS_n \text{ is relayed via } CP_m \text{ and allocated } \omega_n \text{ bandwidth;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\text{(CMRP-MILP)} \qquad \underset{\mathbb{Z},\mathbb{A},\mathbb{B},\mathbb{W}}{maximize} \qquad \mathcal{C} = \sum_{m=1}^{M}\sum_{n=1}^{N} b_{mn} r_{mn} \tag{3.12}$$

subject to:

$$\sum_{m=1}^{M} b_{mn} r_{mn} \geq \rho_n \quad \text{for } \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.13}$$

$$b_{mn} \geq 0 \quad \text{for } \forall m \in \mathbb{N}_{\mathbb{CP}}, \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.14}$$

$$b_{mn} \leq BW a_{mn} \quad \text{for } \forall m \in \mathbb{N}_{\mathbb{CP}}, \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.15}$$

$$b_{mn} \leq BW(1 - a_{mn}) + \omega_n \quad \text{for } \forall m \in \mathbb{N}_{\mathbb{CP}}, \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.16}$$

$$\sum_{m=1}^{M} z_m = K \tag{3.17}$$

$$a_{mn} \leq z_m \quad \text{for } \forall m \in \mathbb{N}_{\mathbb{CP}}, \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.18}$$

$$\sum_{m=1}^{M} a_{mn} = 1 \quad \text{for } \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.19}$$

$$\sum_{n=1}^{N} w_n \leq BW \tag{3.20}$$

$$a_{mn} \in \{0,1\} \quad \text{for } \forall m \in \mathbb{N}_{\mathbb{CP}}, \forall n \in \mathbb{N}_{\mathbb{SS}} \tag{3.21}$$

$$z_m \in \{0,1\} \quad \text{for } \forall m \in \mathbb{N}_{\mathbb{CP}}. \tag{3.22}$$

where $r_{mn} = \min[C(\theta^* \frac{P_{BS}}{d_m^\alpha}), C(\frac{P_{BS}}{d_n^\alpha} + \frac{P_{RS}}{d_{mn}^\alpha} + 2\sqrt{\theta^* \frac{P_{BS} P_{RS}}{(d_n d_{mn})^\alpha}})]$.

Proposition 1: The optimal solution of Problem CMRP is identical with that of Problem CMRP-MILP.

Proof: To prove the proposition, we will show that the objective functions and the constraints of Problem CMRP and CMRP-MILP are equivalent.

Clearly, the objective function of (3.12) and constraint (3.13) in Problem CMRP-MILP are equivalent to those of (3.4) and (3.5) in Problem CMRP, respectively. In addition, the constraints (3.19), (3.20), (3.21) and (3.22), in Problem CMRP-MILP are the same as the constraints (3.6), (3.9), (3.10)and (3.11) in Problem CMRP. To show that the decision variable $(b_{mn})_{M \times N}$ is defined validly in the domain, the constraints (3.14)-(3.16) function as follows. If $a_{mn} = 0$, then $b_{mn} \leq 0$ by (3.15) and $b_{mn} \leq BW + \omega_n$ by (3.16), with (3.14), thus $b_{mn} = 0$. If $a_{mn} = 1$, then $b_{mn} \leq BW$ by (3.15), $b_{mn} \leq \omega_n$ by (3.16), and with (3.14), thus, $0 \leq b_{mn} \leq \omega_n$. Such a relaxation of b_{mn} does not change the solution of the problem. Therefore, the formulations of the two problems are equivalent, and there is no nonlinear constraints and objective function in the Problem CMRP-MILP. In other words, the optimal solution of Problem CMRP equals that of Problem CMRP-MILP. ■

3.2.3 Upper Bound of the Cell Capacity

We derive an upper bound of the cell capacity $\mathcal{C}^{\mathcal{UB}}$. This analytical bound can be used to assist the fast estimation of cell capacity performance given a specific *network configuration*. The *network configuration* here refers to the cell bandwidth and network layout information which consists of the locations of CPs for deploying RSs and SSs. Note that this analytical bound is independent of the number of RSs.

Let $\omega_n = \omega_n^0 + \Delta_n$, where ω_n^0 denotes the minimum bandwidth allocated to SS_n to achieve the rate requirement ρ_n given the associated RS's location information. To minimize the ω_n^0 that is allocated to SS_n while ρ_n is achieved, the rate of SS_n needs to be maximized. Namely, $\omega_n^0 = \min\limits_{\forall m} \frac{\rho_n}{r_{mn}} = \frac{\rho_n}{\max\limits_{\forall m} r_{mn}}$. For SS_n ($\forall n \in \mathbb{N}_{SS}$), $a_{mn} = 1$ when $m = \arg\max\limits_{\forall m} r_{mn}$; otherwise, $a_{mn} = 0$. Then, we can derive the upper bound of cell capacity as follows

$$\begin{aligned}
\mathcal{C} &= \sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn}\omega_n r_{mn} = \sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn}(\omega_n^0 + \Delta_n) r_{mn} \\
&= \sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn}\omega_n^0 r_{mn} + \sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn}\Delta_n r_{mn} \\
&= \sum_{n=1}^{N}\rho_n + \sum_{n=1}^{N}\Delta_n \sum_{m=1}^{M} a_{mn} r_{mn} \\
&\leq \sum_{n=1}^{N}\rho_n + \max\Big(\sum_{n=1}^{N}\Delta_n\big(\sum_{m=1}^{M} a_{mn} r_{mn}\big)\Big) \\
&\leq \sum_{n=1}^{N}\rho_n + (BW - \sum_{n=1}^{N}\omega_n^0)\max_{\forall n,\forall m}(r_{mn})
\end{aligned} \tag{3.23}$$

Therefore, the upper bound of cell capacity is expressed as

$$\mathcal{C}^{\mathcal{UB}} = \sum_{n=1}^{N}\rho_n + (BW - \sum_{n=1}^{N}\frac{\rho_n}{\max_{\forall m} r_{mn}})\max_{\forall n,\forall m}(r_{mn}) = BW \max_{\forall n,\forall m} r_{mn}. \tag{3.24}$$

Remark 1: (3.24) can be employed to evaluate if the capacity requirement can be satisfied with current network configuration. If not, it indicates that the total traffic

demands at all the SSs exceed the capacity that the current network configuration can provide; in other words, upgrading current network configuration with more cell bandwidth and/or more CPs are needed to achieve the design requirement.

Remark 2: Another constraint,

$$\mathcal{C} \leq \mathcal{C}^{\mathcal{UB}}$$

can be added in the formulation of problem CMRP and CMRP-MILP which is capable of speeding up the calculation of solutions due to the reduction of search space.

3.3 Solution Algorithms to RS Placement

3.3.1 An Improved Genetic Algorithm

Based on a global optimization meta-heuristic method, i.e., Generic Algorithm (GA) [85], we propose an approach to solve Problem CMRP. The GA implementation is outlined in Algorithm 1 and elaborated as follows.

• Algorithm Description

As shown in Figure 3.3, the genetic representation (chromosome) of a candidate solution $\mathcal{I}_j^{(i)}$, i.e., the jth individual in the ith generation, is represented by a $1 \times N$ vector of ordered pairs (q_n, w_n), where q_n and w_n denote the associated CP and allocated bandwidth to SS_n, respectively; and $q_n = m$ is equivalent to $a_{mn} = 1$. w_ns are generated, after q_ns, according to the following inequality: $\frac{\rho_n}{r_{q_n,n}} \leq w_n \leq B$, where $r_{q_n,n}$ represents the achievable rate of SS_n relayed with the RS located at CP_{q_n}. The lower bound of ω_n is obtained according to constraint (3.5). w_ns are set within the valid range subjected to the total bandwidth constraint (3.9).

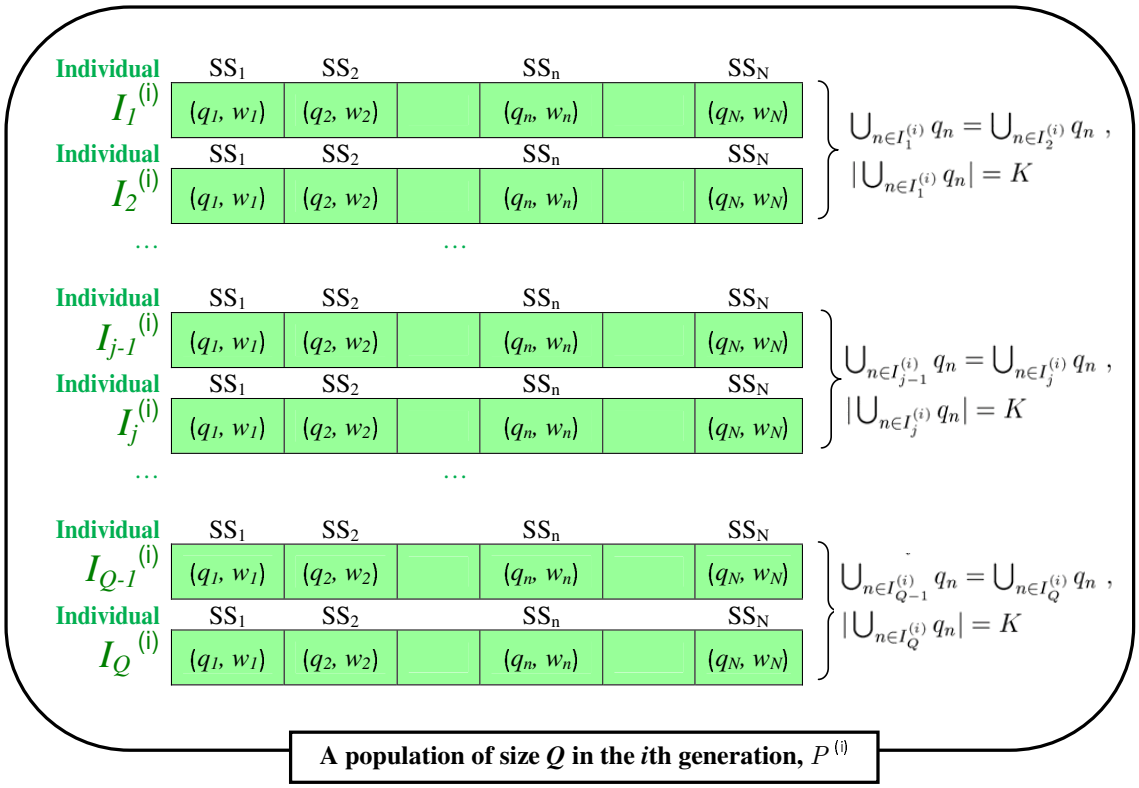


Figure 3.3: Population and individual (chromosome) representation for GA-based heuristic to Problem CMRP

To generate the initial population (*Line 3* of Algorithm 1) of size Q, denoted as $\mathcal{P}^{(1)}$, we randomly select K among M CPs and randomly assign them to q_n in $\mathcal{I}_1^{(1)}$ and $\mathcal{I}_2^{(1)}$. Every two adjacent individuals of population, denoted as $\mathcal{I}_j^{(i)}$ and $\mathcal{I}_{j+1}^{(i)}$, have the same allocated CPs but with different permutation. The same process is repeated for the next pair of individuals of initial population.

GA improves the population through a number of iterations (*Lines 6 to 24*). In each iteration, four operators (selection, replacement, crossover, and mutation) are performed to generate new population.

Algorithm 1: GA implementation for Solving Problem CMRP

Input: **Problem parameters**: $M, N, K, \mathbb{A}_{SS}, \mathbb{A}_{CP}, \mathbb{I}_{CP}, \mathbb{I}_{SS}, \mathbb{P}, P_{BS}, P_{RS}, B, \alpha$;
GA parameters: $Q, N_G, p_c, p_m, \varepsilon, N_T$

begin
  Set up and initialize: Input GA parameters and problem parameters;
  Generate initial population $\mathcal{P}^{(1)}$;
  Find $elite^{(1)}$;
  $i = 2$;
  while $Exit\_flag == 0$ **do**
    Perform *selection* using roulette wheel sampling scheme;
    Replace the worst individual with $elite^{(i-1)}$;
    for $j = 1$ *to* Q **do**
      if $\mathcal{I}_j^{(i)} \neq elite^{(i-1)}$ **then**
        if j *mod 2 ==1* **then**
          while *constraints*(3.8)&(3.9) *are not held* **do**
            Crossover with probability p_c;
          end
        end
        while *constraints*(3.8)&(3.9) *are not held* **do**
          Mutation with probability p_m;
        end
      end
    end
    Find $elite^{(i)}$;
    $\mathcal{P}^{(i+1)} = \mathcal{P}^{(i)}$;
    $Exit\_flag =$ **Check_termin_conditions**;
    $i = i + 1$;
  end
  $\mathcal{C} = \max_{\forall i} f(elite^{(i)})$;
  return $(b_n, \omega_n)_{1\times N}, \mathcal{C}$
end

Selection (Line 7): Selection operates on the *fitness* values of the individuals, $\mathcal{I}_j^{(i)}$, defined by $f(\mathcal{I}_j^{(i)}) = \sum_{n=1}^{N} w_n r_{b_n,n}$. The selection operator is a fitness proportionate selection known as roulette-wheel [85]. It assigns weights to the individuals such that the individual will be more likely to be selected if it has a larger fitness value.

Replacement (Line 8): To preserve the best solution achieved during the evolutions, we substitute the worst individual (with lowest fitness value) with the best one (with highest fitness value). Thus, the individual with the best fitness, i.e., *elite*, always survives to the next generation.

Crossover (Line 13): The crossover operator uses uniform distribution to choose the crossover points. The crossover operates on q_n first according to a crossover probability p_c; then w_ns are calculated. The constraints (3.8) and (3.9) are checked after the crossover, if any of the constraints is violated, the crossover operation is redone.

Mutation (Line 17): The mutation operator uses uniform distribution to choose the mutation elements, and the mutation is conducted according to a probability p_m. The new w_ns are calculated subsequent to the mutation of q_ns. If any of the constraints (3.8) and (3.9) is violated after mutation, the operation is repeated.

The algorithm terminates with the function **Check_termin_conditions** when any of the following conditions holds:

(a) A predefined number of generations N_G are completed;

(b) The resultant capacity of the whole cell is close enough to the upper bound $\mathcal{C}^{\mathcal{UB}}$, i.e., $\mathcal{C}^{\mathcal{UB}} - f \leq \varepsilon$, where ε is a predefined small number;

(c) The improvement of the cell capacity is negligible after a predefined number of iterations N_T.

If none of the above conditions holds *(Line 23)*, $Exit\_flag$ is set as 0 and the algorithm goes back to *Line 7* to conduct next iteration of GA operations; otherwise, the

algorithm returns and output the *elite*, i.e., $(q_n, \omega_n)_{1\times N}$ and the corresponding fitness value, i.e., cell capacity $\mathcal{C}$.

The computation complexity is mainly determined by the population size Q and the number of generations N_G. In each generation, the selection and replacement operations take $O(Q \times N)$ time. Both the crossover and mutation operation take $O(Q \times N^3)$ since they go through each individual in the population and also check the constraints (3.8) and (3.9). Therefore, the total computation complexity for the GA is $O(N_G \times Q \times N^3)$.

• Improved GA for Solving Problem CMRP

To accelerate the convergence speed, we improve the GA from two perspectives: (a) search space reduction, (b) fast convergence of bandwidth allocation.

(a) Search Space Reduction

The search space increases exponentially with the number of CPs. In order to reduce the search space, we exclude the CPs that can not provide a substantial performance gain in terms of achievable rate. As an example, Figure 3.4 shows the corresponding 2D contours of absolute rate gain G_a =0.8 bit/sec/Hz for SS_1, SS_2 and SS_3, respectively. The points outside the contour curves have less rate gain than 0.8 bit/sec/Hz. So the points CP_2, CP_6, CP_7, CP_{11} can be excluded from the set of CPs. The reason is that these points can not provide a large rate gain for SS_1, SS_2, SS_3, and they can not consequently contribute to the overall cell capacity enhancement. The more excluded CPs, the faster convergence of the GA. However, the required gain value cannot increase unlimitedly since there may not exist a solution if G_a is too large.

(b) Fast Convergence of Bandwidth Allocation

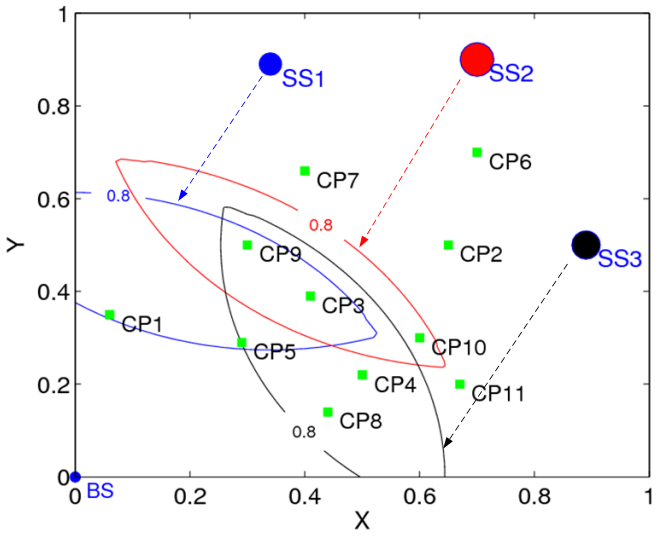


Figure 3.4: Contour of rate gain for corresponding SS, the distance between the BS and SS is normalized, $P_{BS} = 1W$, $P_{RS} = 0.5W$, $\alpha = 2$, $G_a = 0.8bit/sec/Hz$.

Random allocation of both a_{mn}s and ω_ns in each iteration of GA causes the algorithm converges very slowly. It is realized from the (3.4) that $\mathcal{C}$ is a linear function of ω_n when the set of variables a_{mn} are known. Therefore, in each iteration of GA, after a_{mn}s are randomly allocated, instead of random allocation of values to ω_n, we can solve a Linear Program (LP) to obtain the optimal values of ω_n. This strategy combined with elitism causes the algorithm converges much faster. Accordingly, in each iteration, we first randomly allocate K CPs to N SSs so that the constraints (3.6), (3.8), (3.10), and (3.11) are satisfied, then we solve Problem (BW-A), which is expressed as follows:

$$\text{(BW-A)} \qquad \text{maximize} \quad \mathcal{C} = \sum_{m=1}^{M} \sum_{n=1}^{N} a_{mn} \omega_n r_{mn} \tag{3.25}$$

subject to:

$$\sum_{m=1}^{M} a_{mn}\omega_n r_{mn} \geq \rho_n \quad \text{for } \forall n \tag{3.26}$$

$$\sum_{n=1}^{N} w_n \leq B, \tag{3.27}$$

to obtain the optimal set of ω_n for SS_n.

3.3.2 A Fast Heuristic Algorithm

To avoid the exponential computation time and allow a timely adaptation to the traffic demand change in a cell, we further propose a fast heuristic approach to solve Problem CMRP. The pseudcode is shown in Algorithm 3, which is described as follows.

In essence, the process of determining the location of RS for each SS is to select an element from the set of CPs ($\mathbb{N}_{CP}$). After initialization (Line 2), the upperbound of the cell capacity is calculated using (3.24) and stored in $\mathcal{C}^{\mathcal{UB}}$. If $\mathcal{C}^{\mathcal{UB}}$ is less than the total traffic demands of all the SSs (i.e., $\mathcal{C}^{\mathcal{UB}} < \sum_n \rho_n$), it indicates that no solution exists with current network configuration. If $\mathcal{C}^{\mathcal{UB}} \geq \sum_n \rho_n$, then the SSs are sorted in an decreasing order in terms of their traffic demands and indexed in $\mathbb{N}_{SS}$ (Line 8). In Lines 9-16, beginning from SS_1, the SS with the largest traffic demand, all the CPs are enumerated by the function **BestCP4SS** to find the one that can provide the maximized achievable rate (denoted as Q_1), for SS_1. Namely, $Q_1 = \arg \max_{\forall m \in \mathbb{N}_{CP}} r_{m1}$, which is stored in the set of RSs $\mathbb{N}_{RS}$, i.e., the set of selected CPs. Then Q_1 is mapped to the RS-SS association decision variable by the function **QMapA** such that $a_{Q_1,1} = 1$. The same operation repeats for the next SS. The iteration stops when the number of elements in $\mathbb{N}_{RS}$ equals K. For the remaining SSs $\in \mathbb{N}_{SS}$, in Lines 17-22, we still use function **BestCP4SS** to select the best CP in $\mathbb{N}_{RS}$ that offers the maximized achievable rate for the corresponding SS. Similarly, the index of the selected CP is then mapped

Algorithm 2: A Proposed Fast Heuristic Algorithm for Problem CMRP

Input: $M, N, K, \mathbb{N}_{CP}, \mathbb{N}_{SS}, \mathbb{P}, P_{BS}, P_{RS}, BW, \alpha$;

Output: $\mathcal{C}, \mathbb{N}_{RS}, \mathbb{A}, \mathbb{W}$;

begin
  $\mathbb{N}_{RS} = \Phi; \mathbb{A} = zeros([M, N]); n = 1;$// *Initialization.*
  $\mathcal{C}^{UB} \leftarrow$ **Capacity_Estimation**; // *Estimate upperbound of cell capacity given current network configuration.*
  if $\mathcal{C}^{UB} < \sum(\mathbb{P})$ **then**
    return Display ("No solutions exist for current network configuration!");
  end
  else
    Sort($\mathbb{P}$, '*desc*'); Reorder ($\mathbb{N}_{SS}$);
    while $|\mathbb{N}_{RS}| < K$ **do**
      $Q_n \leftarrow$ **BestCP4SS**$(n, \mathbb{N}_{CP})$;
      //*Find the best CP for* SS_n *among* $\mathbb{N}_{CP}$ *such that* $Q_n = \arg \max_{\forall m \in \mathbb{N}_{CP}} r_{mn}$;
      $\mathbb{N}_{RS} \leftarrow Q_n$;
      $\mathbb{A} \leftarrow$ **QMapA** (Q_n);
      //Q_n *is mapped to* $\mathbb{A}$ *such that* $a_{Q_n,n} = 1$;
      $n = n + 1$;
    end
    for $j = n$ *to* N **do**
      $Q_j \leftarrow$ **BestCP4SS**$(j, \mathbb{N}_{RS})$;
      //*Find the best CP for* SS_j *among* $\mathbb{N}_{RS}$ *such that* $Q_j = \arg \max_{\forall m \in \mathbb{N}_{RS}} r_{mj}$;
      $\mathbb{A} \leftarrow$ **QMapA** (Q_j);
      //Q_j *is mapped to* $\mathbb{A}$ *such that* $a_{Q_j,j} = 1$;
    end
    $[\mathcal{C}, (\omega_n)_{1\times N}] \leftarrow$ **SolveLP_BWA**$((a_{mn})_{M\times N})$; // *Allocate bandwidth to each SS given* $\mathbb{A}$ *by solving an LP*;
    return $\mathcal{C}$, $\mathbb{N}_{RS}$, $\mathbb{A} = (a_{mn})_{M\times N}$, $\mathbb{W} = (\omega_n)_{1\times N}$
  end
end

to the RS-SS association decision variable. The constraints (3.6), (3.8), (3.10) and (3.11) in the CMRP formulation are guaranteed to be valid in these steps. Till now, the heuristic has obtained the solution of $\mathbb{N}_{RS}$ and $\mathbb{A}$. Next, in Line 23, the function **SolveLP_BWA** to calculate the optimal values of $\mathbb{W}$ provided $\mathbb{N}_{RS}$ and $\mathbb{A}$ by solving an LP (BW-A) such that the cell capacity is maximized. (BW-A) is defined in previous subsection. Due to the maturity of current technique of solving linear programs, we can efficiently get the results of bandwidth allocation $(\omega_n)_{1\times N}$ and the objective value of cell capacity $\mathcal{C}$ immediately (Line 26), the constraints (3.5) and (3.9) are also ensured.

The analysis of the computational complexity of the above heuristic algorithm is as follows. The computation time for Lines 2-8 is $O(N \log N)$. The process of determining $\mathbb{N}_{SS}$ (Lines 9-16) yields a computation complexity $O(KM)$, and the process of allocating CPs in $\mathbb{N}_{RS}$ to the remaining SSs with smaller traffic demands (Lines 17-22) yields a worst-case complexity $O(N(K+M))$. Line 23 yields the time complexity to solve an LP. Typically, a primal-dual interior point algorithm ensures the $O(\sqrt{N}L)$ iteration polynomial-time computation complexity, where L is the length of the binary representation of all the numbers involved [86]. Therefore, the overall computation complexity is $O(N \log N + KM + N(K+M) + \sqrt{N}L)$.

3.4 Numerical Results

In this section, numerical investigation will be conducted to evaluate the performance of the proposed two approaches, i.e., the GA-based heuristic and the fast algorithm. We will focus on their optimality and computational efficiency against the benchmark, which is the optimal solution obtained by solving the CMRP-MILP using CPLEX. We will also demonstrate the performance benefits of cooperative relaying against those without relaying and traditional multi-hop non-cooperative relaying.

3.4.1 Case Study I: Single Relay

Before going into the scenarios with multiple RSs and SSs, the performance gain in the simplest three-node relay model is evaluated, aiming to provide an insight on the fundamental performance behavior of cooperative relaying. Without loss of generality, the distance between the source and destination is normalized.

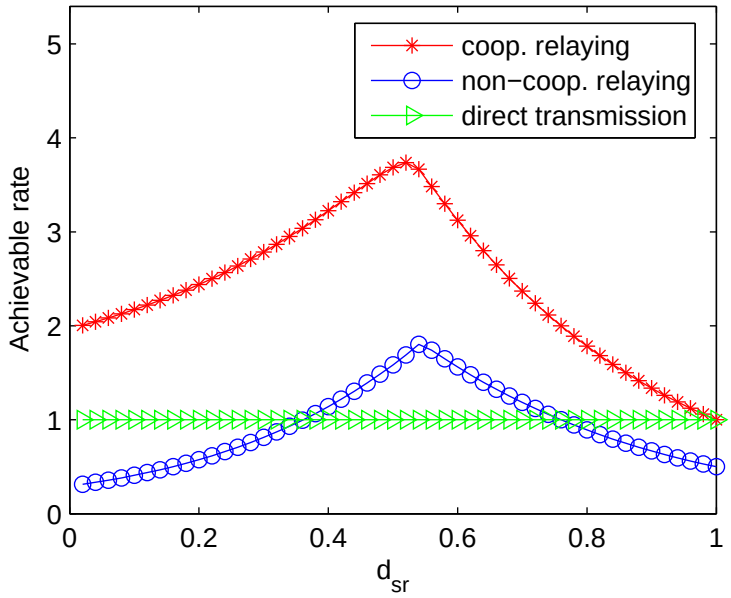


Figure 3.5: Achievable rate (in unit of bits/sec/channel use) comparison in 3-node relay model.

Figure 3.5 shows the achievable rate comparison in the following three mode: (i) cooperative relaying, (ii) non-cooperative relaying, and (iii) direct transmission without relaying. The x-axis represents the normalized distance between the relay and the source. It is observed that the traditional 2-hop non-cooperative transmission rate is not always larger than the direct transmission rate, although a relay node is deployed and leads to extra CAPEX of RS equipment. In other words, if cooperative transmission

is not employed, the relay will not definitely increase the rate since it introduces an extra hop which results in the total limited resource is shared by 2 links. However, the destination's achievable rate with cooperative relaying is always larger than that with the non-cooperative relaying and direct transmission. Figure 3.5 validates that by placing a relay at the same location, cooperative relaying outperforms traditional non-cooperative relaying greatly in terms of rate improvement or equivalently coverage extension.

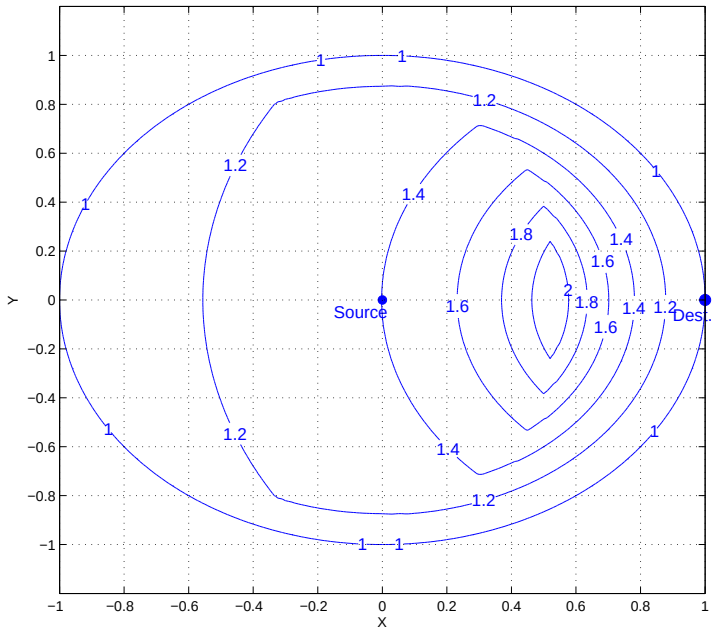


Figure 3.6: 2D contour of relative rate gain.

Figure 3.6 shows 2D contour of relative rate gain. The coordinates in x-axis and y-axis represent the normalized position of a relay. The number shown in each contour line is the corresponding relative rate gain when the relay is placed at the point on the contour line. The relative rate gain is defined as $G_r = \frac{R_{relay}}{R_0}$, where R_0 is the

direct transmission rate. This figure can provide some insight on the destination's rate improvement affected by the relay location in a planar.

3.4.2 Case Study II: Multiple RSs and SSs

To simulate practical CMRP problem, three cases are investigated, where an IEEE 802.16j cell with an OFDM interface is assumed in the case studies. The transmit power of the BS and RS are set to 1w and 0.5w, respectively. The path loss exponent α is set to 3. The bandwidth allocated to the cell is 20MHz. The thermal noise is assumed to be constant at each subcarrier with equal subcarrier bandwidth. To simulate the presence of buildings, trees and other obstructions in practical network environments, the shadowing effects typically result in 5-10 dB losses are also taken into account in the simulations.

The case studies are conducted in the following steps. (1) The CMRP-MILP formulation firstly solved using CPLEX 10.0 [84] to obtain the optimal solution, which is taken as a benchmark to evaluate the optimality gap by our proposed heuristic counterparts; (2) The convergence of the GA and improved GA solutions are evaluated; (3) The optimality and computation time for the improved-GA and fast heuristic algorithm are compared; (4) The performance gain of cooperative relaying is demonstrated against the traditional non-cooperative relaying and direct transmissions.

For step (1), three scenarios are investigated and compared in the simulation. Figure 3.7 (a), (b) and (c) illustrate the network layout of scenario (I), (II) and (III), respectively. The coordinate of each SS in the cell is normalized, and the amount of traffic demand is proportional to the radius of the circle representing the SS. The problem size, the number of constraints, variables, and nonzero elements of the constraint matrix, the average computation time, memory consumption, and optimality gap of

the scenarios are shown in Table 3.2. It can be seen that the complexity of solving the CMRP-MILP grows dramatically as the network size increases, which indicates the difficulty in achieving timely adaptation for practical large-scale network planning.

For step (2), the GA parameters we use in the simulation are listed in Table 3.3. Figure 3.8 demonstrates that both GA and improved GA heuristics are capable of approaching to the optimality. In addition, the convergence of the GA and improved GA is compared. The parameter G_a in the improved GA can be manipulated such that the solution in the initial generation yielded by improved GA is better than that without search space reduction. On the other hand, the average slope of the curve for the improved GA is larger than that of the GA without fast bandwidth allocation. Hence, the results show that the two methods employed in the improved GA can effectively accelerate the convergence.

For step (3), we compare the improved GA and the fast heuristic algorithm against the benchmark on the optimality gap as well as the computation time. The results are shown in Table 3.4. Figure 3.9 shows the cell capacity with different number of RSs for placement in Scenario (I). From Table 3.4 and Figure 3.9, it can be seen that the improved GA can provide a better solution than the fast algorithm, while the fast algorithm has better computation efficiency than the improved GA, with only slight degradation on the capacity performance. If the number of RSs is close to the number of best CPs and a rapid deployment for timely response is required, the fast algorithm is a priori solution to Problem CMRP; otherwise, the improved GA is a better solution due to the better optimality achieved.

For step (4), we present the performance gain of RS placement. The "*best CPs*" shown in Figure 3.7(a)-(c) represent the ones with maximum achievable rate for the corresponding SSs and are determined by an exhaustive search among all the CPs.

The resultant cell capacity can achieve the capacity upper bound when each *best CP* is place an RS.

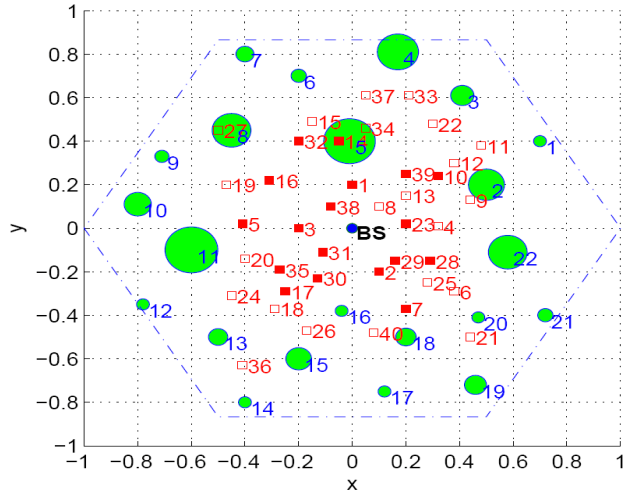
Figure 3.10 compares each SS's achievable rate with (i) direct transmission, (ii) cooperative relaying, and (iii) non-cooperative relaying in each scenario, respectively. We can observe that the data rate of each SS with relaying has a significant increase compared with that of without relaying. For instance, in Scenario (I), with cooperative relaying, the achievable rate for each SS increases from 46.27% to 113.87% over the direct transmission rate; while with non-cooperative relaying, the achievable rate for each SS increases from 3.02% to 101.37%. Figure 3.10 further demonstrates that cooperative relaying can achieve better rate than non-cooperative relaying. And the end-user rate enhancement with optimal RS placement contributes substantially to the overall system performance gain in terms of the cell capacity enhancement and coverage extension.

3.5 Summary

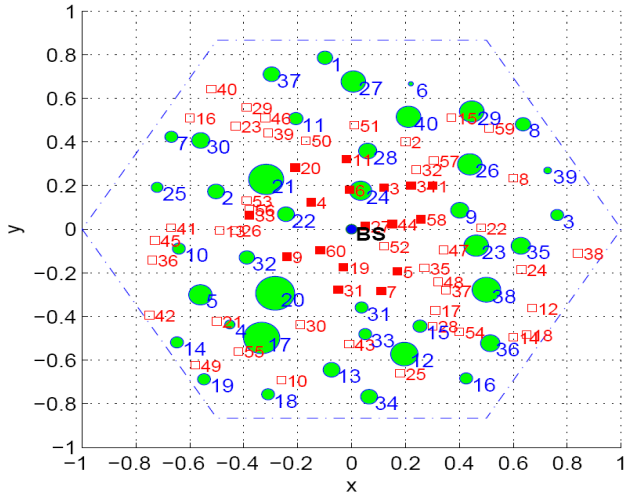
In this chapter, we have investigated the RS location planning in broadband wireless access networks, where cooperative relaying in employed. An optimization framework of joint optimal RS location and bandwidth allocation has been proposed. The design objectives are to maximize the overall cell capacity as well as meet the traffic load demand of each SS. We have firstly defined the CMRP problem, which is an MINLP. An improved GA and a fast heuristic algorithm have been developed to reduce the exponential computation time. In addition, an upper bound on the cell capacity has been derived to assist the fast estimation of system performance. Numerical analysis has demonstrated the effectiveness and efficiency of the proposed solutions and the significant performance gains due to RS placement under cooperative relaying.

Table 3.1: Definitions of Important Symbols for Problem CMRP

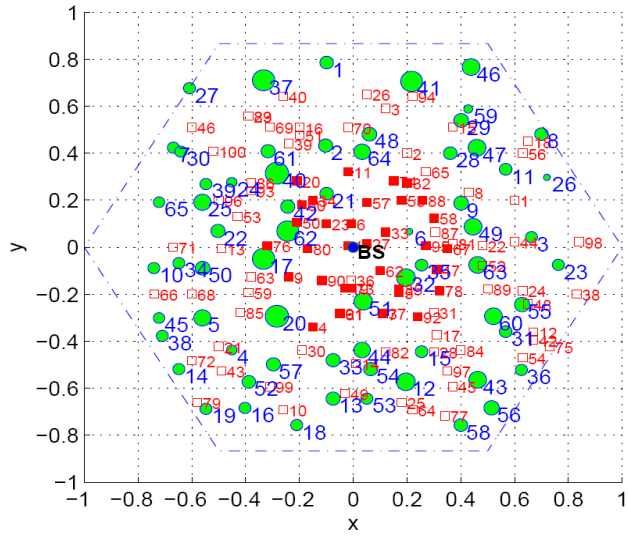
Symbol	Definition
$\mathbb{A}$	The RS-SS incidence matrix (decision variable), $\mathbb{A} = (a_{mn})_{M\times N}$.
$\mathbb{W}$	The bandwidth-allocation vector (decision variable), $\mathbb{W} = (\omega_n)_{1\times N}$.
$\mathbb{Z}$	The RS location incidence vector (decision variable), $\mathbb{Z} = (z_m)_{1\times M}$.
K	The number of RSs to be deployed within the cell.
$\mathbb{N}_{\mathbb{SS}}$	The set of SSs, $\lvert\mathbb{N}_{\mathbb{SS}}\rvert = N$.
$\mathbb{N}_{\mathbb{CP}}$	The set of CPs, $\lvert\mathbb{N}_{\mathbb{CP}}\rvert = M$.
$\mathbb{N}_{\mathbb{RS}}$	The set of RSs, $\lvert\mathbb{N}_{\mathbb{RS}}\rvert = K$.
ρ_n	The minimum traffic demand for SS_n.
$\mathbb{P}$	The set of minimum traffic demand for SSs, $\mathbb{P} = \{\rho_n, \forall n \in \mathbb{N}_{\mathbb{SS}}\}$.
d_{ij}	The distance between node i and node j.
d_m	The distance between the BS and CP_m.
d_n	The distance between the BS and SS_n.
α	Attenuation factor.
θ	The power allocation ratio at source between the relay path and direct transmission path.
θ^*	The optimal power allocation ratio at source.
P_{BS}	The transmission power of BS.
P_{RS}	The transmission power of RS.
BW	The upper bound of radio bandwidth allocated to the cell.
$\mathcal{C}$	The cell capacity.
$\mathcal{C}^{\mathcal{UB}}$	The upper bound of the cell capacity.
Q	The population size in GA.
N_G	The number of generations in GA.
p_c	The crossover probability in GA.
p_m	The mutation probability in GA.
N_T	The predefined number of iterations in GA.
ε	The predefined small number of threshold for stopping criteria (b) in GA.
$\mathcal{I}_j^{(i)}$	The jth individual in the ith generation in GA.
$\mathcal{P}^{(i)}$	The population in the ith generation in GA.
G_a	The absolute rate gain.
G_r	The relative rate gain.
R_0	The direct transmission rate without relaying.



(a) Scenario (I)



(b) Scenario (II)



(c) Scenario (III)

SS CP Best CP

Figure 3.7: Illustration of the cell layout.

Table 3.2: Problem Size, Computation Time, Amount of Memory Occupied for Cplex Solving the CMRP-MILP Formulation

Scenario	Num. of Nodes		Constraints	Variables	Nonzeros	Memory	Opt. Gap
	SS	CP					
(I)	22	40	2687	1822	7982	58.01 MB	3.57%
(II)	40	60	7283	4900	21700	565.88 MB	4.14%
(III)	65	100	19633	13165	58665	7372.23 MB	6.20%

Table 3.3: GA Parameter Setting

GA parameter	Q	N_G	p_c	p_m	ε	N_T
Setting	10000	50000	0.8	0.01	0.001	5000

Table 3.4: Comparison in Computation Time and Optimality Gap against the Benchmark of CPLEX

Scenario	Optimality Gap against Benchmark		Computation Time (sec)	
	Improved GA	Fast algorithm	Improved GA	Fast algorithm
(I)	1.57%	5.71%	1134	1.197
(II)	2.25%	8.22%	8856	3.422
(III)	4.86%	9.52%	98246	9.418

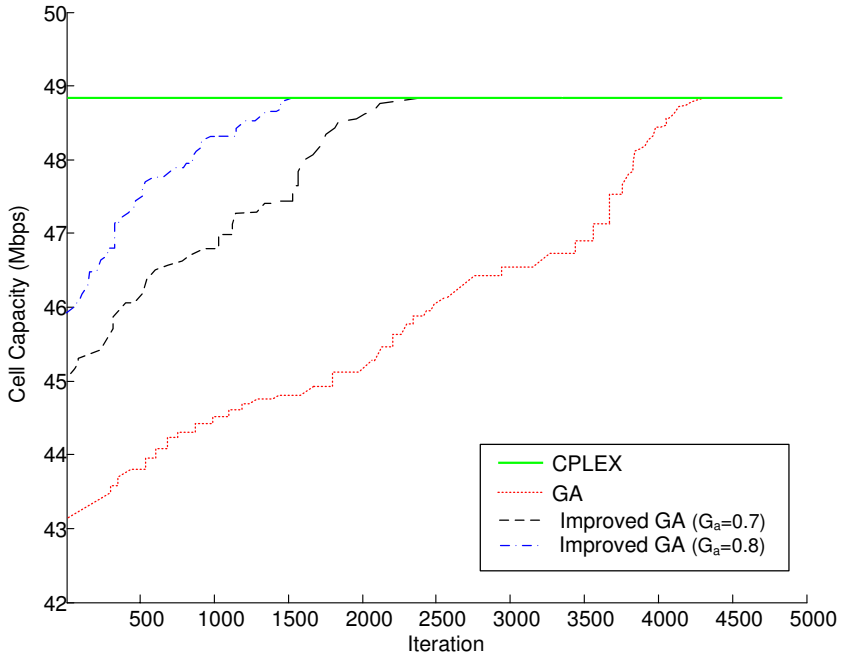


Figure 3.8: Convergence comparison of GA and improved GA.

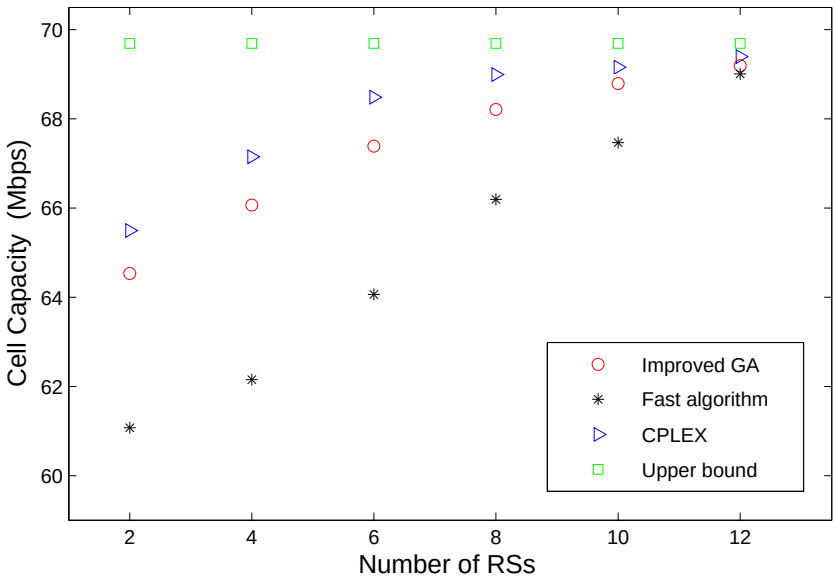
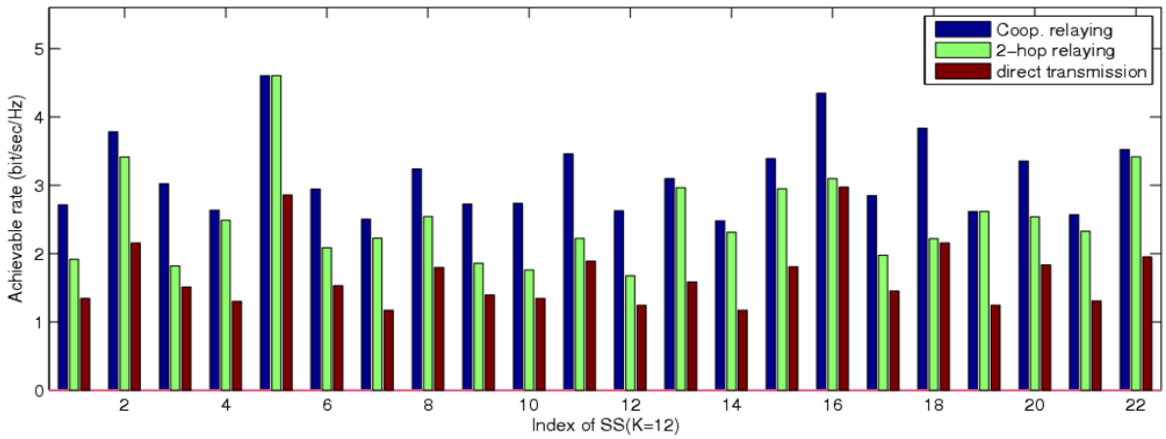
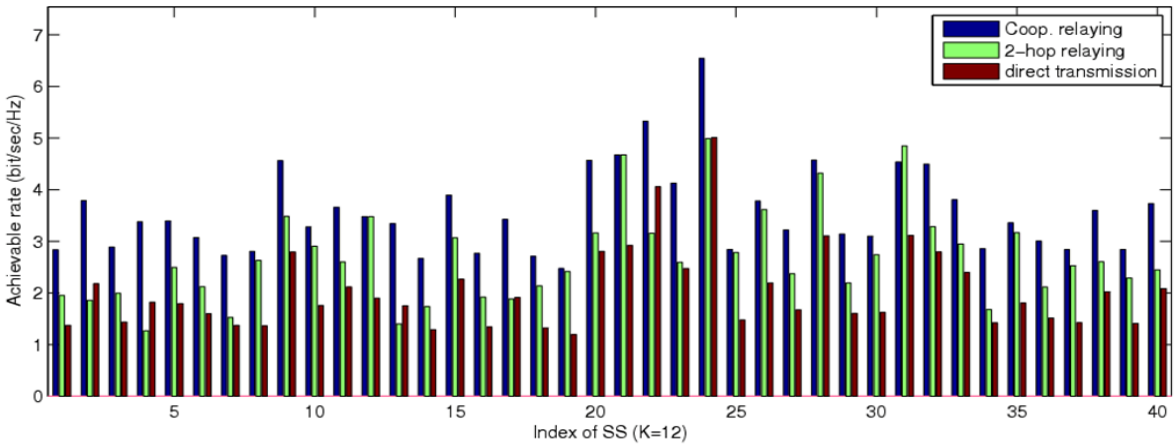


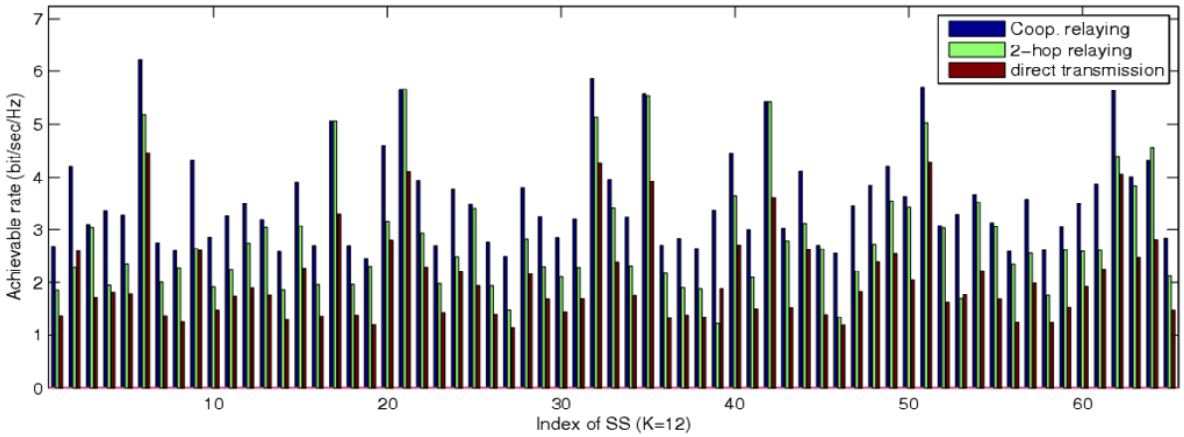
Figure 3.9: Cell capacity vs. number of RSs in Scenario (I).



(a) Scenario I



(b) Scenario II



(c) Scenario III

Figure 3.10: Achievable rate for each SS where 12 RSs are deployed.

Chapter 4

Location Planning under Multi-level Cooperative Relaying

In this chapter, we address the issue of **D**imensioning and **L**ocation **P**lanning (also called **Problem DLP**) in a more complex multi-hop broadband wireless environment [87]. Extending the study of Chapter 3 with the assumption of 2-hop relaying, in this chapter, we will explore Problem DLP under *multi-level cooperative relaying (CR)*.

The *multi-level CR* in this research refers to a multi-hop transmission that both the source and internal relays cooperatively transmit to a destination [7, 28]. It is different from the traditional *multi-hop non-cooperative transmission (NCT)*, in which the received data are simply forwarded at each internal node to the next-hop node[1]. In other words, a multi-level CR transmission scheme exploits the concurrent transmission of signals to the destination, instead of taking the signals sent by different relay nodes as *interference* with each other as that in the traditional multi-hop NCT.

[1] In this study, we use the terminology "*multi-level CR*" to distinguish from the traditional "*multi-hop NCT* ".

4.1 Modeling of Multi-level Cooperative Relaying

In the fundamental multi-level relay channel model, a source (node 0) communicates with a destination (node G) through multiple relays (node $1, 2, \ldots, G-1$). The achievable rate for a fixed destination can be expressed as [9]:

$$R = \min_{1 \le j \le G} [C(\frac{1}{N_0} \sum_{k=1}^{j} (\sum_{i=0}^{k-1} \sqrt{P_{ik}} d_{ij}^{-\alpha})^2)] \tag{4.1}$$

where j is the level (stage) of intermediate relay node, G is the maximal level of destination node, α is the path loss exponent, P_{ik} is the transmit power of node i to node k, N_0 is the power of the background noise, $C(.)$ is the *Shannon function* such that $C(x) = \frac{1}{2}\log(1+x)$ for $x \ge 0$, with d_{ij} as the distance between node i and j.

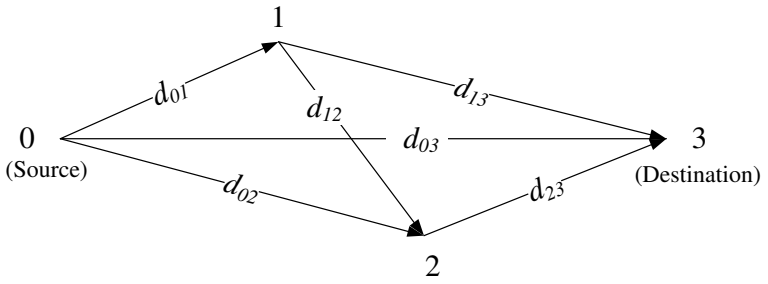


Figure 4.1: An illustration of 3-level relaying.

As an example shown in Figure 4.1, there are two internal relays, i.e., $G = 3$. The achievable rate for the destination (node 3) is

$$R_3 = \min[R_3^{(1)}, R_3^{(2)}, R_3^{(3)}] \tag{4.2}$$

where

$$R_3^{(1)} = C(\frac{P_{01}}{N_0 d_{01}^{2\alpha}}),$$
$$R_3^{(2)} = C[\frac{1}{N_0}[\frac{P_{01}}{d_{02}^{2\alpha}} + (\frac{\sqrt{P_{02}}}{d_{02}^{\alpha}} + \frac{\sqrt{P_{12}}}{d_{12}^{\alpha}})^2]],$$
$$R_3^{(3)} = C[\frac{1}{N_0}[\frac{P_{01}}{d_{03}^{2\alpha}} + (\frac{\sqrt{P_{02}}}{d_{03}^{\alpha}} + \frac{\sqrt{P_{12}}}{d_{13}^{\alpha}})^2 + (\frac{\sqrt{P_{03}}}{d_{03}^{\alpha}} + \frac{\sqrt{P_{13}}}{d_{13}^{\alpha}} + \frac{\sqrt{P_{23}}}{d_{23}^{\alpha}})^2]].$$

For simplicity, we assume that all the nodes use the same power density for transmission, i.e., $P_{ik} = P_0$ for $\forall i, k$.

4.2 Problem Formulation

We consider a practical deployment scenario in a multi-hop broadband wireless access network. Similar to the three-tier network architecture and radio propagation model in *Chapter 3*, the locations and minimal rate requirements for a set of dispersed SSs, the locations of CPs for deploying RSs, and the transmission power of the BS and RSs are known as prior information before network planning.

Definition 4.1. Problem DLP. The goals of Dimensioning and Location Planning (DLP) problem under multi-level CR are to

(1) obtain the minimal number of RSs to be deployed,

(2) determine the optimal locations of the RSs,

(3) assign serving RS(s) to each SS,

(4) and determine the relay sequence of the serving RS(s) for each SS.

The notations used in the problem are listed in Table 5.1. In the following, we will provide the mathematical formulation to capture the nature of the problem, and characterize the behavior and constraints under multi-level CR.

Table 4.1: Definitions of Important Symbols for Problem DLP

Symbol	Definition
N_{CP}	The set of CPs for RSs within the cell, $\|N_{CP}\| = M$.
N_{SS}	The set of SS within the cell, $\|N_{SS}\| = N$.
Ω	The set of all the nodes within the cell, $\Omega = \{BS\} \cup N_{CP} \cup N_{SS}$, $\|\Omega\| = 1 + M + N$.
ρ_n	The minimal rate requirement for SS_n.
P_0	The transmit power of an RS.
N_0	The average thermal noise power in cell.
α	The path loss exponent.
ML	The maximal allowable number of intermediate RSs in a cell.
$\overrightarrow{\mathscr{F}^n}$	The virtue directed flow for SS_n.
q_A^D	The relay level of node A for destination D.
$\mathcal{C}$	The objective value of the problem DLP.
$\mathbb{Z}$	The CP-RS location incidence vector.
$\mathbb{U}$	The CP-SS association incidence matrix.
$\mathbb{F}$	The virtue flow matrix.
$\mathbb{Q}$	The relaying node level (stage) matrix.
$\mathbb{H}$	The CP-relay-level incidence matrix.
$\mathscr{F}_j^{n,t}$	The virtue flow (relay sequence) of SS_n with t relays in column j in SS_n's sub-table.
$r_j^{n,t}$	The achievable rate of SS_n with t relays in column j in SS_n's sub-table.
j	The column index in an SS's sub-table.
J_n	The number of columns in SS_n's sub-table after truncation.
$\xi(n,m)$	The contributive index of $\text{CP}_m(\in N_{CP})$ to $\text{SS}_n(\in N_{SS})$.
$\Xi(m)$	The set contributive index of $\text{CP}_m(\in N_{CP})$ to N_{SS}.
$\mathcal{C}^{LB}$	The lower bound of the objective value.

Let $\overrightarrow{G} = (\Omega, \overrightarrow{E})$ represent the directed graph, which consists of a node set $\Omega = \{BS\} \bigcup N_{CP} \bigcup N_{SS}$ and a directed edge set $\overrightarrow{E}$. N_{CP} and N_{SS} are the set of CPs and SSs, respectively. To formulate the problem, we start by defining the decision variables for RS location ($\mathbb{Z}$) and allocation ($\mathbb{U}$), respectively. $\mathbb{Z} = (z_m)_{1\times M}$ is a CP-RS incidence vector (*location variable*) where $z_m = 1$ if CP_m is selected for placing an RS; otherwise, $z_m = 0$. $\mathbb{U} = (u_{mn})_{M\times N}$ is a CP-SS incidence matrix (*CP allocation variable*) where $u_{mn} = 1$ if CP_m is allocated to serving SS_n; otherwise, $u_{mn} = 0$.

Since the DLP problem targets at minimizing the total number of RSs, denoted by $\mathcal{C}$, the objective function can be expressed as

$$\text{minimize} \quad \mathcal{C} = \sum_{m\in N_{CP}} z_m \tag{4.3}$$

The RS location and allocation variables (i.e., $\mathbb{Z}$ and $\mathbb{U}$) are subject to the following constraints.

$$u_{mn} \leq z_m, \quad \forall m \in N_{CP}, n \in N_{SS} \tag{4.4}$$

$$\sum_{m\in N_{CP}} u_{mn} \leq ML, \quad \forall m \in N_{CP}, n \in N_{SS} \tag{4.5}$$

Constraint (4.4) ensures that if CP_m is associated with SS_n, an RS must be placed at CP_m. Constraint (4.5) makes a limit for the maximal number of internal relays (i.e., associated CPs) for an SS.

The complexity of the formulation arises at representing the achievable rate for each SS in practical deployment scenario. It is because each CP is still undetermined whether to place an RS and whether to be allocated to an SS. Moreover, even the allocation of RSs to an SS has been determined, their exact sequence in a relay route is still undecided, since there are totally $n!$ permutations of possible sequences for n RSs. To represent the relay sequence, we adopt a concept of *virtue flow*, which is defined as

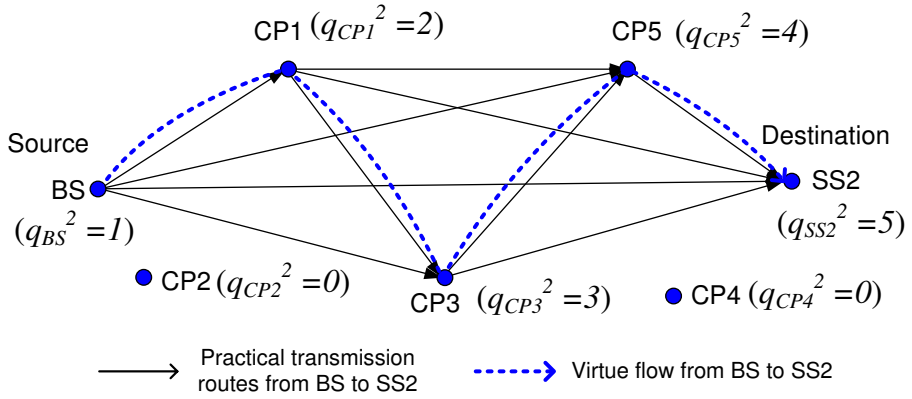


Figure 4.2: An illustration of practical relay routes and virtue flow (BS→ SS_2).

the route with the maximal number of hops among all the practical transmission routes from BS to SS_n. The virtue flow (or relay sequence) is denoted by $\overrightarrow{\mathscr{F}^n}$. For example, in Figure 4.2, suppose CP_1, CP_3 and CP_5 are allocated to provision transmissions to SS_2, and CP_2 and CP_4 are not. The virtue flow $\overrightarrow{\mathscr{F}^2}$ is illustrated with a blue dotted line. Namely, $\overrightarrow{\mathscr{F}^2}$: {BS, CP_1, CP_3, CP_5, SS_2}, for simplicity, $\overrightarrow{\mathscr{F}^2}$: {CP_1, CP_3, CP_5}.

Then, we define a virtue flow matrix $\mathbb{F} = (f_{ij}^n)_{(M+N+1)\times(M+N+1)\times N}$, where $f_{ij}^n = 1$ if both node i and node j are in $\overrightarrow{\mathscr{F}^n}$, and node i is on the upstream of node j (for $i \in BS \bigcup N_{CP}$, $j \in N_{CP} \bigcup SS_n$, $n \in N_{SS}$); otherwise, $f_{ij}^n = 0$. Figure 4.3 shows the matrix of virtue flow $\overrightarrow{\mathscr{F}^2}$ from BS to SS_2.

Due to the large number of zero elements in the sparse matrix (f_{ij}^n), we compress the matrix to reduce the number of variables. In the compressed matrix of $\mathbb{F}$, the rows and columns that are all zeros in all situations are removed. Specifically, $\mathbb{F} = (f_{ij}^n)_{(M+1)\times(M+1)\times N}$, where $i \in \{BS\} \bigcup N_{CP}$, $j \in N_{CP} \bigcup SS_n$ and $n \in N_{SS}$. In Figure 4.4, the compressed matrix of (f_{ij}^2) is shown. For each SS, the virtue flow matrix is

$(f_{ij}^2)=$

	BS	CP1	CP2	CP3	CP4	CP5	SS1	SS2	...	SSn
BS	0	1	0	0	0	0	0	0	...	0
CP1	0	0	0	1	0	0	0	0	...	0
CP2	0	0	0	0	0	0	0	0	...	0
CP3	0	0	0	0	0	1	0	0	...	0
CP4	0	0	0	0	0	0	0	0	...	0
CP5	0	0	0	0	0	0	0	1	...	0
SS1	0	0	0	0	0	0	0	0	...	0
SS2	0	0	0	0	0	0	0	0	...	0
...	...		...		...		...	...	...	...
SSn	0	0	0	0	0	0	0	0	...	0

Figure 4.3: The virtue flow matrix $\mathbb{F}$ for SS_2.

$(f_{ij}^2)=$

	CP1	CP2	CP3	CP4	CP5	SS2
BS	1	0	0	0	0	0
CP1	0	0	1	0	0	0
CP2	0	0	0	0	0	0
CP3	0	0	0	0	1	0
CP4	0	0	0	0	0	0
CP5	0	0	0	0	0	1

Figure 4.4: The compressed virtue flow matrix $\mathbb{F}$ for SS_2.

subjected to the following constraints:

$$f_{ij}^n + f_{ji}^n \leq 1, \quad \forall i, j \in N_{CP}, n \in N_{SS} \tag{4.6}$$

$$\sum_{m \in N_{CP}} f_{BS,m}^n = 1, \quad \forall n \in N_{SS} \tag{4.7}$$

$$\sum_{i \in \Omega} f_{i,SS_n}^n = 1, \quad \forall n \in N_{SS} \tag{4.8}$$

$$\sum_{i \in N_{CP} \cup \{BS\}} f_{mi}^n = u_{mn}, \quad \forall m \in N_{CP}, n \in N_{SS} \tag{4.9}$$

$$\sum_{i \in N_{CP} \cup BS} f_{im}^n = u_{mn}, \quad \forall m \in N_{CP}, n \in N_{SS} \tag{4.10}$$

$$\sum_{j_1 \in N_{CP}} f_{j_1 m}^n + \sum_{j_2 \in N_{CP}} f_{m j_2}^n = 2u_{mn}, \quad \forall m \in N_{CP}, n \in N_{SS} \tag{4.11}$$

Constraint (4.6) stipulates that the virtue flow is directed where $f_{ij}^n = 1$ if node i is on the upstream of j in $\overrightarrow{\mathscr{F}^n}$. Constraints (4.7) and (4.8) ensure that the BS and each terminal node (SS_n) in $\overrightarrow{\mathscr{F}^n}$ has only one downstream node and one upstream node, respectively. Constraints (4.9) and (4.10) ensure that each CP in $\overrightarrow{\mathscr{F}^n}$ also has exactly one upstream node as well as one downstream node, if the CP is associated with SS_n (i.e., the CP is in $\overrightarrow{\mathscr{F}^n}$). Constraint (4.11) formulates the flow conservation property; i.e., the number of upstream and downstream nodes of a CP is equal to two if the CP is in $\overrightarrow{\mathscr{F}^n}$.

To represent the position of a node in a relay sequence, we now define the *relay level* of a node as follows.

Definition 4.2. Relay level. The *relay level* of node A for destination D (denoted by q_A^D) is defined as $q_A^D = \sum_{A\in\Omega} \deg^-(A)+1$, where $\sum_{A\in\Omega} \deg^-(A)$ represents the *indegree* of node A in graph G.

Note that $q_A^D = 0$ indicates that node A is not associated with destination D. In Figure 4.2, the relay levels of BS, CP_1 to CP_5 are 1, 2, 0, 3, 0 and 4, respectively.

We define a node relay level matrix $\mathbb{Q} = (q_i^n)_{|\Omega|\times N}$, where q_i^n is the relay level of node $i \in \Omega$. According to the constructed directed virtue flow, the relay level of the BS should always be 1. The relay level of node j increases by one, if both i and j are in $\overrightarrow{\mathscr{F}}^n$ and the upstream node of j is i ; otherwise, the relay level is set to 0. The following constraints (4.12)-(4.14) stipulate the above definition of relay level.

$$q_{BS}^n = 1, \quad n \in N_{SS} \tag{4.12}$$

$$q_j^n - q_i^n \geq f_{ij}^n - \beta(1 - f_{ij}^n), \forall i \in N_{CP} \cup BS, j \in N_{CP} \cup SS_n, n \in N_{SS} \tag{4.13}$$

$$q_i^n \geq 0, i \in \Omega, n \in N_{SS} \tag{4.14}$$

where $\beta = |\Omega|$. To establish a connection between the relay level of a CP and Eq.

(4.1), a new set of variable $\mathbb{H}$ is needed. $\mathbb{H} = (H^n_{hi})_{ML\times M\times N}$ is a CP-relay-level incidence matrix such that $H^n_{hi} = 1$ if node $i \in N_{CP}$ is in $\overrightarrow{\mathscr{F}^n}$ and i's relay level is $h \in \{2, \ldots, ML+1\}$; otherwise, $H^n_{hi} = 0$. The following constraints (4.15) and (4.16) stipulate the definition of $\mathbb{H}$. Constraint (4.17) ensures that there is at most one node at each level in $\overrightarrow{\mathscr{F}}^n$.

$$H^n_{hi} \geq \gamma + \gamma(q^n_i - h), \forall i \in N_{CP}, n \in N_{SS}, h \in \{2, \ldots, ML+1\} \tag{4.15}$$

$$1 - H^n_{hi} \geq \gamma + \gamma q^n_i, \quad \forall i \in N_{CP}, n \in N_{SS}, h \in \{2, \ldots, ML+1\} \tag{4.16}$$

$$\sum_{i\in N_{CP}} H^n_{hi} \leq 1, \qquad \forall n \in N_{SS}, h \in \{2, \ldots, ML+1\} \tag{4.17}$$

where $\gamma = \frac{1}{|\Omega|^2}$. To ensure the throughput of each SS is larger than the minimal rate requirement, the constraints can be expressed as follows.

$$\log[1 + (\frac{1}{N_0}\sum_{k=2}^{h_2}(\sum_{h_1=1}^{k}\sqrt{P_0}(D^n_{h_1h_2})^{-\alpha})^2)] \geq \rho_n, h_2 = 2 \ldots ML+1, \forall n \in N_{SS} \tag{4.18}$$

$$D^n_{h_1h_2} = \sqrt{(\sum_{l\in\Omega} H^n_{h_1l}x_l - \sum_{l\in\Omega} H^n_{h_2l}x_l)^2 + (\sum_{l\in\Omega} H^n_{h_1l}y_l - \sum_{l\in\Omega} H^n_{h_2l}y_l)^2}, \tag{4.19}$$

$$\forall 2 \leq h_1 < h_2 \leq ML+1, \forall n \in N_{SS}$$

$$D^n_{1,h_1} \leq \frac{D^n_{1,h_2}}{\varepsilon + \sum_{l\in\Omega} H_{h_2l}}, \forall 2 \leq h_1 < h_2 \leq ML+1, n \in N_{SS} \tag{4.20}$$

where ε is a very small number. Constraint (4.20) ensures that the relay sequence of the internal relays (the selected CPs) is the one with maximal achievable rate among all the permutations of the nodes.

In summary, the objective function (4.3) and the constraints (4.4) to (4.20) form the mathematical formulation for problem DLP. Due to the nonlinearity of constraints (4.18), (4.19) and (4.20), problem DLP is a nonlinear integer program, which cannot be solved by any systematic method. In the next section, we present an effective algorithm to solve problem DLP by exploiting the nature of the problem.

4.3 An Optimal Dimensioning and Location Planning Algorithm

4.3.1 Algorithm Description

By further observing the formulation of problem DLP, we find that the key obstacle in solving this problem lies in the huge search space with myriad combinations of possible solutions. In particular, searching the best relay sequence (virtue flow) for each SS, i.e., variable $\mathbb{Q}$, in the huge search space is extremely difficult. However, once the optimal relay sequence is determined, the achievable rate for each SS is yielded, and then $\mathbb{Z}$, $\mathbb{U}$, and $\mathcal{C}$ can be yielded accordingly.

To this end, we propose a two-phase solution procedure: i) Feasible-Relay-Sequence (FRS) table setup phase; ii) minimum cost relay planning phase, which is shown in Algorithm 3. The core idea is as follows. First, we reduce the search space such that only feasible candidate relay sequences are left. Then, in Phase II, we iteratively remove the least contributive CP till the best relay sequence for each SS is found.

- Phase I: Feasible-Relay-Sequence (FRS) Table Setup Phase

The search space is defined in the FRS table, which is the basis for deciding relay locations in the next phase. As an example, assume $ML = 3$, the FRS table is shown in Figure 4.5. Each SS has its own sub-table. The sub-table consists of two rows: a row of achievable rates, which are sorted in non-increasing order, and a row of relay sequences, which one-to-one map the rates in the same column in the first row. $r_j^{n,t}$ and $\mathscr{F}_j^{n,t}$ represent the achievable rate of $\text{SS}_n \in N_{SS}$ via t relays and the corresponding relay sequence (i.e., virtue flow) for SS_n, respectively. j is the column index.

Algorithm 3: **A Proposed Two-Phase Algorithm for Problem DLP**

Input: $N_{CP}, N_{SS}, (\rho_n)_{1\times N}, M, N, c, P_0, \alpha, ML$;
Output: $\mathcal{C}, \mathbb{Z}, \mathbb{U}, \mathbb{Q}$;
Initialization: Input the given cell layout and system parameters ;
Phase I: Feasible-Relay-Sequence (FRS) Table Setup
for $\forall\ SS_n \in N_{SS}$ **do**
  Calculate $r_1^{n,1}$, the achievable rate of SS_n with direct transmission;
  for $t = 2$ *to* ML **do**
    for $j = 1$ *to* ${}^{M}C_t$ **do**
      Calculate $r_j^{n,t}$, the maximal achievable rate of SS_n for the j^{th} combination of t selected CPs in N_{CP};
      Store the corresponding relay sequence in $\mathscr{F}_j^{n,t}$ to achieve the rate $r_j^{n,t}$;
    end
    Sort $r_j^{n,t}$s in non-increasing order and write entries in the first row of achievable rate in SS_n's sub-table;
    Map the corresponding $\mathscr{F}_j^{n,t}$ and write relay sequence entries in SS_n's sub-table;
  end
  if $\rho_n > r_1^{n,ML}$ **then**
    Print{"No solution exists with current network configuration!"}; **return**;
  else
    Determine t_n, the required number of relays, for SS_n by Eq. (4.21);
    Truncate the SS_n's sub-table such that only the columns satisfying $r_j^{n,t} \geq \rho_n$ and $t = t_n$ are remaining;
    Calculate the lower bound of the objective value $C^{LB} = \max_{\forall n \in N_{SS}} J_n$;
  end
end
Phase II: Minimum Cost Relay Planning
Set the initial set of N_{RS} as the union of all the CPs in the FRS table, i.e., $N_{RS} = \bigcup_{n\in N_{SS}} \bigcup_{j\in\{1\ldots J_n\}} \mathscr{F}_j^{n,t_n}$;
Calculate contributive index $\xi(n,m)$ and set contributive index $\Xi(m)$ for $\forall m \in N_{RS}, \forall n \in N_{SS}$;
/* FRS_Filter procedure:*/
repeat
  Compute $\Delta = \max_{\forall m\in N_{RS}}(\Xi(m)) - \min_{\forall m\in N_{RS}}(\Xi(m))$;
  Find CP_m with minimal set contributive index, i.e., $m = \arg\min_{\forall m\in N_{RS}}(\Xi(m))$;
  if $\exists n$ *such that* $m \in \bigcap_{j\in\{1\ldots J_n\}} \mathscr{F}_j^{n,t_n}$ **then**
    Set CP_m as optimal, i.e., $\mathbb{Z} \leftarrow m$;
  else
    Update N_{RS} by removing CP_m, i.e., $N_{RS} \leftarrow N_{RS}/\{m\}$;
    Update SS_n's sub-table for $\forall n \in N_{SS}$ by removing the column that contains CP_m in SS_n's sub-table; $J_n \leftarrow J_n - \xi(n,m)$;
    Re-calculate $\xi(n,m)$ and $\Xi(m)$ for $\forall m \in N_{RS}$, $\forall n \in N_{SS}$;
  end
until *(*$\Delta == 0$*) or (*$|\mathscr{F}^n| == 1$ *for* $\forall n \in N_{SS}$*) or (*$|N_{RS}| == C^{LB}$*)*
Set the optimal relay sequences of RSs ($\mathbb{Q}$) as the first column ($\mathscr{F}_1^{n,t_n}$) in each SS's updated sub-table;
Determine the final locations ($\mathbb{Z}$), the allocation ($\mathbb{U}$) to each SS, and total number of RSs ($\mathcal{C}$) according to $\mathbb{Q}$;
return;

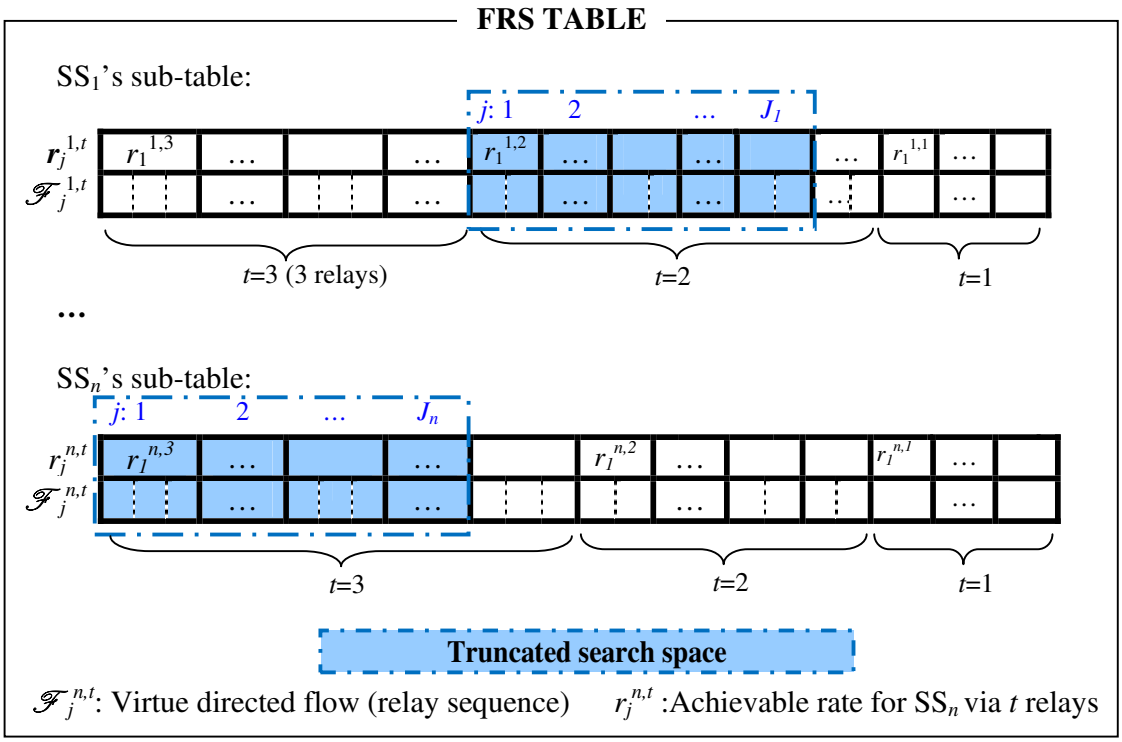


Figure 4.5: An illustration of FRS table ($ML = 3$).

Phase I focuses on setting up such a FRS table. To construct it, for each SS, $r_j^{n,t}$ is calculated with Eq. (4.1) by enumerating all the possible combinations of t relays (CPs) in N_{CP}, where $t = \{1 \dots ML\}$. Note that there are $t!$ number of permutations for a specific set of t relays, which lead to $t!$ achievable rates. Thus, the maximal one is set as $r_j^{n,t}$ and the corresponding relay sequence is stored in $\mathscr{F}_j^{n,t}$. Then, in Line 10, we sort all the $r_j^{n,t}$s in non-increasing order, and map the corresponding $\mathscr{F}_j^{n,t}$. The entries of $r_j^{n,t}$s and $\mathscr{F}_j^{n,t}$s are then written in the FRS table. In Line 13, we check if there exists a feasible solution with current network configuration (the set of CPs). More specifically, we compare ρ_n and $r_1^{n,ML}$, which is the maximum possible rate for SS_n. If $\rho_n > r_1^{n,ML}$, it indicates there exist no feasible solutions with current network configuration (the set of CPs), so more CPs with better locations are needed in N_{CP}. Here, "better" CP means that if an RS is placed at the CP, a satisfying rate gain can be achieved. If $\rho_n \leq r_1^{n,ML}$ for $\forall n$, there must exist a feasible solution. By exploiting the monotonicity of the achievable rate row in SS_n's sub-table, we can use the following criteria to determine the required number of relays (denote as t_n) for SS_n to achieve its minimum rate requirement.

$$t_n = \begin{cases} 1, & \text{for } r_1^{n,0} < \rho_n \leq r_1^{n,1}; \\ 2, & \text{for } r_1^{n,1} < \rho_n \leq r_1^{n,2}; \\ \dots & \\ ML, & \text{for } r_1^{n,ML-1} < \rho_n \leq r_1^{n,ML}. \end{cases} \tag{4.21}$$

$r_1^{n,0}$ represents the direct transmission rate of SS_n without relaying. Afterwards, we truncate the search space in each SS's sub-table such that only the columns satisfying $r_j^{n,t} \geq \rho_n$ and $t = t_n$ are remaining. J_n is the total number of columns in SS_n's sub-table after truncation. In Figure 4.5, only the shadowed columns are left in each SS's

sub-table. The truncated FRS table is ready for the operations in Phase II.

- Phase II: Minimum Cost Relay Planning Phase

In this phase, we propose a FRS_Filter method to find the minimum cost placement, which is a procedure of iteratively removing the least contributive CP in N_{SS}. We now first introduce two new concepts called *contributive index* and *set contributive index* as follow.

Definition 4.3. Contributive Index. The *contributive index* of $\text{CP}_m(\in N_{CP})$ to $\text{SS}_n(\in N_{SS})$ (denoted as $\xi(n,m)$) is defined as the total number of CP_m in the relay sequence row of SS_n's sub-table.

Definition 4.4. Set Contributive Index. The *set contributive index* of $\text{CP}_m(\in N_{CP})$ to the set N_{SS} (denoted as $\Xi(m)$) is defined as the total number of CP_m in all the relay sequence rows in the FRS table, i.e., $\Xi(m) = \sum_{n \in N_{SS}} \xi(n,m)$.

Remark. $\Xi(m)$ represents the contributive index of a specific CP_m to the whole set of N_{SS}. The smaller the value of $\Xi(m)$, the less chance that CP_m is selected to the final optimal set of RSs. Based on this concept, in Phase II, the minimal cost placement will be realized through the process of FRS_Filter.

In Line 22, we first get the initial set of RSs (N_{RS}), which is the union of all CPs in the FRS table. Then, the $\xi(n,m)$s and $\Xi(m)$s for $\forall m \in N_{RS}, \forall n \in N_{SS}$ are calculated according to Definition 4.3 and 4.4. In Lines 24-35, the procedure of FRS_Filter is executed. Suppose $\text{CP}_m(\in N_{RS})$ is the one with minimal set contributive index. We then check if it appears in each column of relay sequence in the SS_n's sub-table. If yes, it indicates that CP_m must be selected to place an RS and it can not be removed, i.e., CP_m is an optimal location for RS; otherwise, we can update N_{RS} by removing CP_m. Then all the SSs' sub-tables should also be updated by removing the whole column

that contains CP_m. Afterwards, all the $\xi(n,m)$s and $\Xi(m)$s are re-calculated so as to conduct next loop of FRS_Filter operations. The loop terminates when any of the following is true:

(1) Δ equals zero, where $\Delta = \max_{\forall m \in N_{RS}}(\Xi(m)) - \min_{\forall m \in N_{RS}}(\Xi(m))$. Δ represents the difference of set contributive index of the remaining CPs in N_{RS}. $\Delta == 0$ indicates that all the CPs left in the FRS table have an equal chance to be finally selected, namely, none is worse than the others. Thus, the current N_{RS} is the optimal, and $\mathbb{Z}$ can be determined accordingly.

(2) $|\mathscr{F}^n|$ equals one for $\forall n \in N_{SS}$. It indicates that there is only one column remaining in each SS's sub-table, in which the left relay sequence is definitely optimal.

(3) $|N_{RS}|$ equals C^{LB}. It indicates that the lower bound of the objective value has been achieved, so the remaining relay sequences are definitely optimal.

Then we set the optimal relay sequence of RSs ($\mathbb{Q}$) to each SS as the first column in the updated SS_n's sub-table; the objective value $\mathcal{C}$, the final selected locations for RSs $\mathbb{Z}$ and relay allocation to each SS $\mathbb{U}$ can also be determined based on the optimal relay sequences $\mathbb{Q}$.

In summary, FRS_Filter guarantees the convergence by repeatedly discarding the one with least possibility to be optimum in the search space defined in FRS table until all the remaining are optimal. Besides assuring convergence, we claim that the proposed algorithm can yield an optimal solution, which will be proved in detail later. The optimality of this algorithm can also be validated in numerical studies, in which we compare the results obtained by our proposed algorithm with the optimal solution obtained by exhaustive search.

4.3.2 Computational Complexity Analysis

The analysis of the computational complexity of the above two-phase algorithm in the worst case is as follows. In Phase I, the complexity of determining the FRS table is $O(N^2 + N \cdot ML + \sum_{t=1}^{ML}({}^{M}C_t \log({}^{M}C_t)))$.[2] In Phase II, the calculation of the contributive index needs to go through every SS and CP, which yields complexity of $O(NM)$. For the procedure of FRS_Filter, the operation in Lines 25-34 in the worst case needs $M-1$ iterations, and recalculation of the contributive index is needed within each iteration, thus FRS_Filter yields the time complexity $O(NM^2)$. After the completion of FRS_Filter loops, the complexity of determining decision variables is $O(N)$. Therefore, the overall computation complexity of the algorithm is $O(N^2 + N \cdot ML + \sum_{t=1}^{ML}({}^{M}C_t \log({}^{M}C_t)) + NM^2)$. Since the FRS table can be constructed in advance, the running time of our algorithm mainly depends on the Phase II's process, which is polynomial time complexity.

4.3.3 Optimality Proof

In this subsection, we give an optimality proof of the proposed 2-phase algorithm, that is, upon the termination of the proposed algorithm, the solution (i.e., the objective value and the relay sequence for each SS) is optimal. The key idea in the proof is based on contradiction, which exploits two important properties of the FRS table as follows:

Property 1. *Sorted lists of the achievable rate in the FRS table.*

Specifically, after Phase I, in the achievable rate row of each SS's sub-table, all the $r_j^{n,t}$s are sorted in non-increasing order. In other words,

$$r_1^{n,ML} \geq r_j^{n,(\cdot)} \geq r_{j+1}^{n,(\cdot)} > r_1^{n,0}, \text{ for } \forall j, \forall n \in N_{SS}$$

[2] ${}^{a}C_b$ represents the number of combinations of b elements in a set of a elements.

$$r_{(\cdot)}^{n,a} > r_1^{n,b} \geq r_{(\cdot)}^{n,b}, \text{ for } \forall a > b \in \{1 \dots ML\}, \forall n \in N_{SS}$$

where $(\cdot)$ represents any feasible element that is within the domain. $r_1^{n,0}$ represents the achievable rate for SS_n with direct transmission. The sorted property makes it easier to determine the required number of relays for each SS through Eq. (4.21). By taking advantage of the monotonicity of the achievable rate in the FRS table, we reduce the optimization problem to a feasibility problem first. In other words, we discard the infeasible solutions in the search space in Phase I of the proposed algorithm.

Property 2. *The search space defined in the FRS table does not exclude the optimal solution after truncation.*

Compared with the exhaustive search, the search space is $[{}^{M}C_{ML} + {}^{M}C_{ML-1} + \ldots + {}^{M}C_1]^N$. The search space grows exponentially and is prohibitive to get an optimal result for large-size problem. In contrast, the search space defined in the FRS table has been reduced significantly such that the solutions that are infeasible to be optimal have been removed, namely, all the entries that satisfy $r^{n,t} < \rho_n$ and $t \neq t_n$ for $\forall n \in N_{SS}, t \in \{1 \dots ML\}$ have been removed. For this property, we need to prove that, at the end of Phase I of the algorithm, the truncated search space in the FRS table does not loose the optimal solution as follows.

Proof. The proof is based on contradiction. Suppose $\mathscr{F}^{n,t_n^*}$ is the optimal relay sequence for $\text{SS}_n \in N_{SS}$, but the column that contains $\mathscr{F}^{n,t_n^*}$ is abandoned in SS_n's sub-table. $\mathscr{F}^{n,t_n}$ is the final relay sequence obtained by our proposed algorithm. We assume all the other SSs' sub-tables do not loose optimal ones. Note that we only need to prove the contradiction in the case that a single optimal entry is lost, because if there are more than one entries excluded in the FRS table, it can be deduced from the single-lose case straightforwardly. According to the

mathematical formulation of Problem DLP, the final objective value is yielded by $\mathcal{C} = \sum_{m \in N_{CP}} z_m = |\mathbb{Z}|$. In other words, the total number of RSs can be determined by the final relay sequence for each SS in essence, i.e., $\mathcal{C} = |\bigcup_{n=1}^{N} \mathscr{F}^{n,t_n}|$.

If the assumption is true, then $\mathcal{C}^* < \mathcal{C}$. It indicates that $t_n^* \leq t_n - 1$. By PROPERTY 1, we have

$$r^{n,t_n^*} \leq r_1^{n,t_n-1} < r_1^{n,t_n}. \tag{4.22}$$

According to the criteria of Eq. (4.21) in our algorithm,

$$r^{n,t_n-1} < \rho_n \leq r_1^{n,t_n}. \tag{4.23}$$

By transitivity of inequalities of (4.22) and (4.23), we have

$$r^{n,t_n^*} < \rho_n. \tag{4.24}$$

(4.24) indicates that the achievable rate by $\mathscr{F}^{n,t_n^*}$ is less than the minimal rate requirement, which is contradicted with constraint (4.18) defined in the mathematical formulation of the problem DLP. The contradiction results from the assumption of the optimality of t_n^*. This completes the proof for PROPERTY 2. ■

We are now ready to prove the following theorem, which is the main result of this subsection.

Theorem 1. *The obtained solution by the proposed 2-phase algorithm is optimal.*

Proof. (By contradiction) Denote $(\mathbb{Z}, \mathbb{U}, \mathbb{Q})$ the final solution obtained by our proposed 2-phase algorithm with the objective value being $\mathcal{C}$. Assume there exists a better solution $(\widetilde{\mathbb{Z}}, \widetilde{\mathbb{U}}, \widetilde{\mathbb{Q}})$ with the objective value $\widetilde{\mathcal{C}} < \mathcal{C}$. Based on PROPERTY 2, the relay sequences of $\widetilde{\mathbb{Q}}$ and $\mathbb{Q}$ must both exist in the FRS table.

According to the mathematical formulation of problem DLP, $\mathcal{C} = \sum_{m \in N_{CP}} z_m = |\mathbb{Z}| = |\bigcup_{n=1}^{N} \mathscr{F}^{n,t_n}|$, thus we have $|\widetilde{\mathbb{Z}}| \leq \mathbb{Z} - 1$. Note that we only need to prove the contradiction when $|\widetilde{\mathbb{Z}}| = \mathbb{Z} - 1$ because if $|\widetilde{\mathbb{Z}}| = \mathbb{Z} - i$ $(i \geq 2)$, then it becomes even worse and can be deduced from the $i = 1$ case.

Let $\widetilde{\mathbb{Z}} = \mathbb{Z}/\{k\}$, i.e., k is not in $\widetilde{\mathbb{Z}}$ and is not an optimal relay location. This will be a basis for the contradiction in the following proof for THEOREM 1. Besides k, all the other elements in $\widetilde{\mathbb{Z}}$ and $\mathbb{Z}$ are the same. By assumption, $\exists m \in \mathbb{Z}$ such that k could be replaced by m. In other words, k should be removed due to its less contributive than other CPs in $\mathbb{Z}$, i.e., $k = \arg\min_{\forall m \in N_{RS}} \Xi(m)$ at some iteration.

According to the mechanism of FRS_Filter in our algorithm, if $k = \arg\min_{\forall m \in N_{RS}} \Xi(m)$, only the following two possibilities exist.

(i) Another CP (say $s \neq k$) with the same set contributive index exists in the same iteration ($s = \arg\min_{\forall m \in N_{RS}} \Xi(m)$). s is chosen to be removed from N_{RS}. As a result, s must not be in $\mathbb{Z}$. This is a valid operation of FRS_Filter.

(ii) While k is chosen and ready to be removed, however, we find that k appears in every relay sequence of some SS's sub-table, i.e., $\exists n \in N_{SS}$ such that $k \in \bigcap_{j \in \{1 \ldots J_n\}} \mathscr{F}^{n,t_n}$. It implies that if k is removed, then all the columns containing $\mathscr{F}^n$ in the SS's sub-table will be removed, and the SS's sub-table will become empty. It is invalid and forbidden for the FRS_Filter process. Thus, k can not be replaced by any $m (m \in \mathbb{Z}/\{k\})$ and must be an optimal relay location. This conflicts with the aforementioned assumption that "*k is not an optimal relay location*". Thus, our solution is optimal and no better solutions than ours exist. The proof is complete. ■

4.4 Numerical Results

4.4.1 Case Study I: 1-D Multi-level Cooperative Relaying

Before going into the scenarios with multiple destinations (SSs), the performance gain in a simple rectilinear multi-level CR model with a single destination is investigated. The goal is to provide an insight on the fundamental performance behavior of multi-level CR. We assume the distance between any two neighbor nodes is a unit distance as shown in Figure 4.6. The transmit power of the BS and RSs are set as 10w and 2w, respectively. The path loss exponent α is 4. It is assumed all the receivers are subject to Additive White Gaussian Noise (AWGN) with zero mean and unit variance. Figure 4.7 shows the achievable rate at destination versus the increase of distance between source and destination. It is observed that the destination's rate can be improved by placing relays between the source and the destination. Furthermore, with the same number of internal relays deployed at the same positions between the source and the destination, multi-level CR significantly outperforms traditional multi-hop NCT.

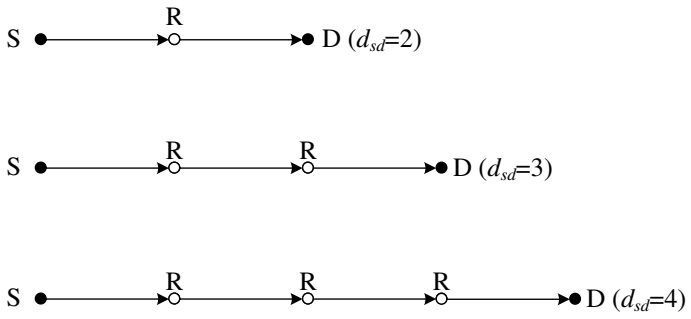


Figure 4.6: The rectilinear multi-level CR model.

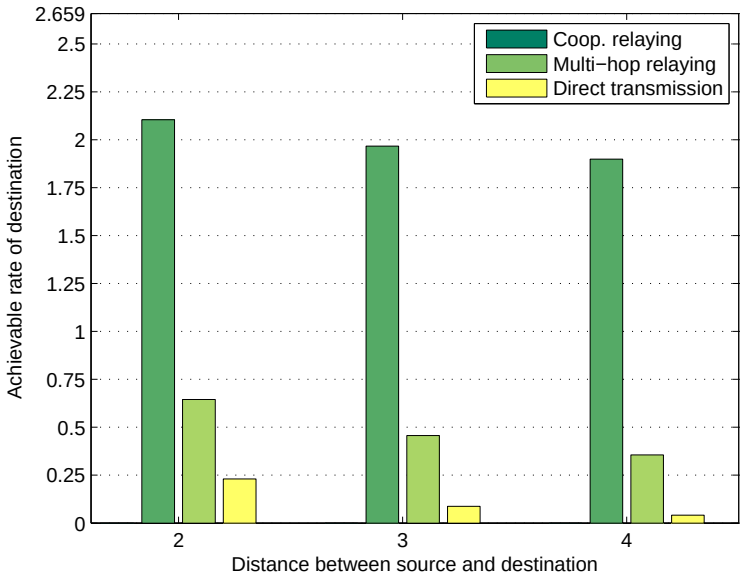


Figure 4.7: The achievable rate of destination vs. distance between source and destination.

4.4.2 Case Study II: 2-D Multi-level Cooperative Relaying

We now consider a practical network scenario where an IEEE 802.16 WirelessMAN air interface is assumed. Without loss of generality, the BS is located at the origin point (0,0). Figure 4.8 shows one of the network layouts considered in the case studies, in which the normalized coordinate of each SS and CP is illustrated and the amount of minimal rate requirement for each SS is proportional to its radius. ML is set to 2. Each node transmits with a unit power density. The path loss exponent α is set to 3. To simulate the presence of buildings, trees and other obstructions in practical network environments, the shadowing effects typically result in 5-10 dB losses are also taken into account in the simulations.

To verify the optimality of our proposed two-phase algorithm, we compare the

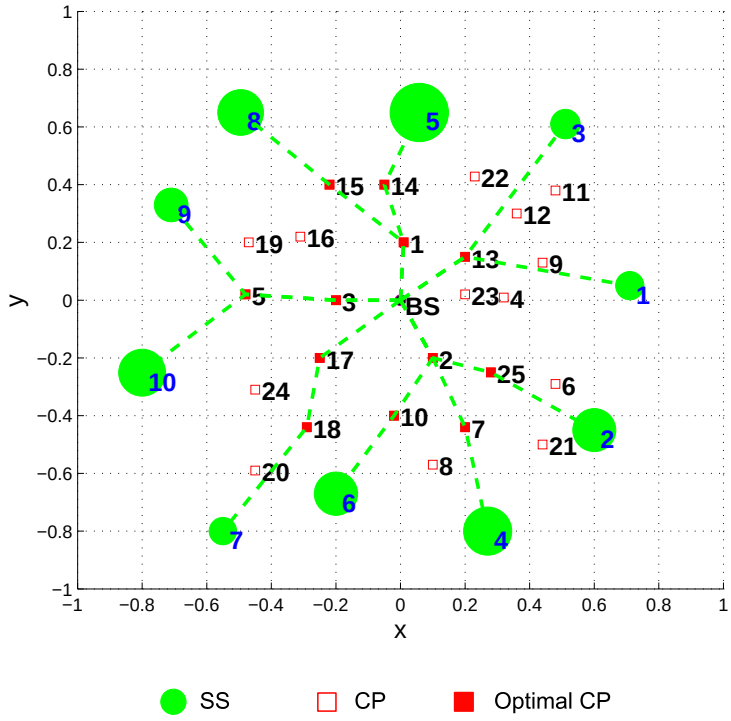


Figure 4.8: The optimal results obtained by proposed 2-phase algorithm.

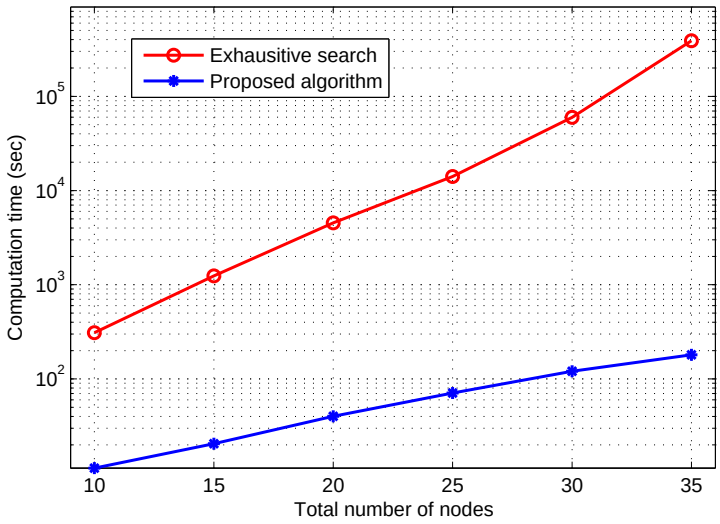


Figure 4.9: The comparison of computation time of proposed algorithm against exhaustive search.

results with the optimal one (obtained by an exhaustive search). We find that the final objective values of our proposed algorithm are the same with the optimum ones for all the cases we studied. Figure 4.8 also shows the resulting configurations of the RS placement and the optimal relay sequences (virtue flows), which are represented with the dotted line from BS to each SS.

The proposed algorithm demonstrates excellent computation efficiency as shown in Figure 4.9. The computation time of our proposed algorithm increases linearly with the problem size (i.e., the total number of SSs and CPs), while the computation time shows an exponential growth with the exhaustive search. Therefore, our proposed algorithm shows a good scalability in practical large-scale wireless network design.

Table 4.2 compares the objective value (i.e., the number of RSs) with multi-level CR against traditional multi-hop NCT. It is observed that the number of RSs to be deployed can be substantially reduced (37.5% $\sim$ 54.6%) when multi-level CR is em-

Table 4.2: Objective Value Comparison

	Number of SSs	4	5	6	7	8	9	10
Objective value	multi-level CR	5	7	8	9	9	10	12
	multi-hop NCT	8	12	15	17	19	22	24
	Saved Cost	37.5 %	41.7 %	46.7 %	47.1 %	52.6 %	54.6 %	50.0 %

ployed. Therefore, by incorporating advanced cooperative transmission technology, a salient economic benefit in terms of RS deployment cost reduction can be achieved. On the other hand, given the same RS placement, a significant improvement on end-user and system throughput can be achieved with cooperative transmission, which just meets the ever-increasing demands for the future residential and business premises.

4.5 Summary

In this chapter, we have conducted a comprehensive study on the issue of dimensioning and location planning in multi-hop wireless networks under multi-level CR. The tasks of RS placement, relay allocation, and determination of relay sequence have been considered together and formulated into a unified optimization framework. To avoid the intractable computation complexity in solving the obtained nonlinear formulation, a two-phase algorithm has been developed. Simulation and case studies have been conducted, and the results have demonstrated the effectiveness and efficiency of the proposed algorithm. The RS deployment cost can be reduce by 37.5% $\sim$ 54.6% with multi-level CR compared to that with traditional multi-hop NCT.

Chapter 5

Location Planning in Integrated PON-WCN

Chapters 3 and *4* have studied the RS placement problem in single-cell scenario. This chapter explores a more complex integrated PON-WCN architecture. It is different from traditional hybrid network architecture, in which the wired and wireless networks are simply interconnected through a common standard interface (e.g., Ethernet) [21] [22]. A key technology of *multi-cell CT* is incorporated due to the aforementioned great performance benefits. Again, we will focus on the problem of network **D**imensioning and **S**ite **P**lanning, which is also called Problem **DSP-PW** [20]. In this research, mobile coverage, BS-user association, wireless bandwidth and power breakdown assignment are also jointly considered in a unified optimization framework. First, let's take a closer look at the system model of integrated PON-WCN.

5.1 System Model of Integrated PON-WCN

In the integrated architecture, the WCN serves as a front-end access network and the PON is desirable to be the backhaul. As shown in Figure 5.1, in the architecture demonstrated in [19], an Optical Network Unit (ONU) in a PON and a wireless BS can be integrated into a single device box, called ONU-BS. Such an integration not only decreases hardware cost, but also achieves an overlaid control plane since the ONU-BS possesses full information on bandwidth request, allocation, and packet scheduling of both the ONU and the wireless BS, optimal mechanisms can be adopted for bandwidth requests in the upstream direction of the PON network, and bandwidth allocation and packet scheduling in the downstream direction of the wireless network.

Downlink inter-cell collaborative transmissions can be initiated through dispersed ONU-BSs in provisioning data to an SS. The prerequisite of implementing CT through space-time/frequency coding technique is a strong support of a central control unit. Thanks to the central coordination and dynamic control capabilities of an Optical Line Terminal (OLT) in the PON, the integrated architecture just satisfies this need. Our goals are to exploit both the performance benefits due to multi-cell CT as well as the interplay advantages due to optical-wireless network integration.

The point-to-multipoint (PMP) topology and downlink broadcast nature of PON also precipitate the implementation of multi-cell CT. Specifically, the OLT broadcasts downstream data to ONU-BSs, and each cooperative ONU-BS transmits a codeword symbol which is orthogonal (or quasi-orthogonal) to all others under the control of the OLT. An SS thus receives and aggregates multiple versions of the same data from each cooperative ONU-BS through independent wireless channels. By way of collaboration among ONU-BSs, virtually formed antenna arrays can be taken as a macro distributed multiple-input single-output (MISO) system. Such an MISO system brings significant

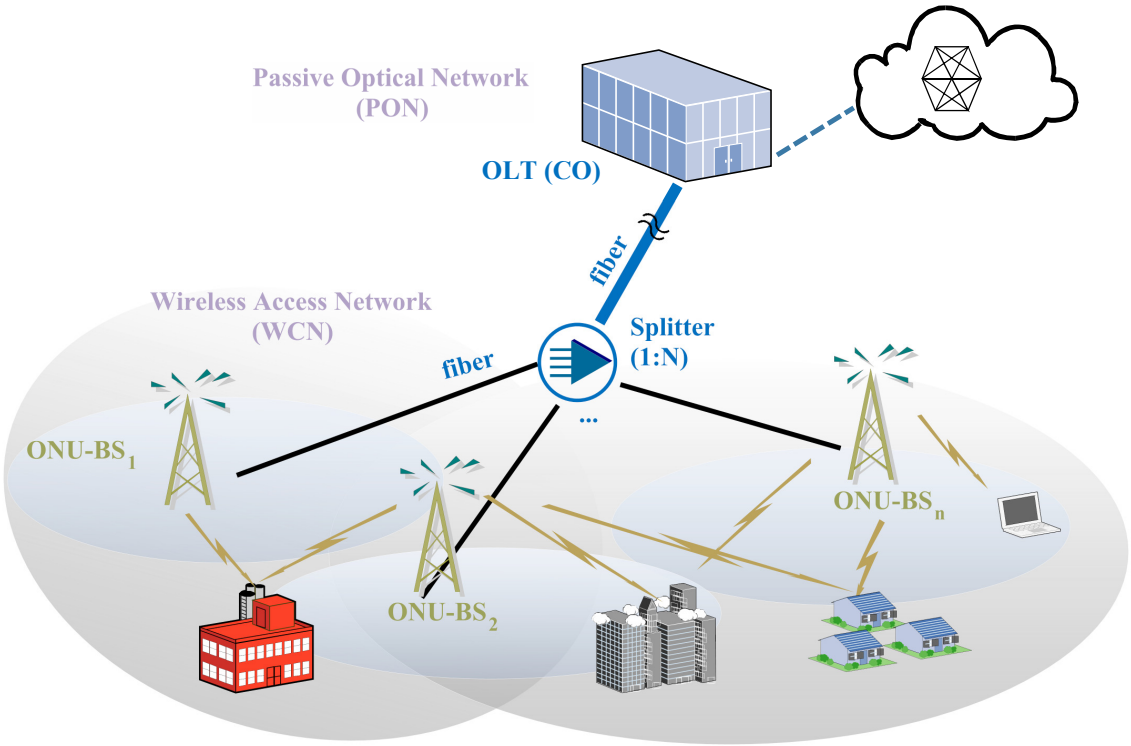


Figure 5.1: Integrated PON-WCN model

benefits from cooperative diversity gain [24] and/or spatial multiplexing gain [88] over conventional non-CT mode.

5.1.1 Main Entities in Integrated PON-WCN

A general network architecture for the integrated PON-WCN is shown in Figure 5.1, which consists of the following key entities: an OLT, optical splitters, fibers, and ONU-BSs. As the root node of the tree structure, the OLT functions as a central controller/coordinator of the whole integrated network. It has powerful computational capability and intelligence. The OLT connects to the global Internet via metropolitan-area network backbone such as Synchronous Optical Networks (SONET), which is

beyond the access network. To extend from the OLT with fibers, a passive optical splitter fans out to multiple optical fibers connected to the ONU-BSs. In addition to the capabilities of conventional BSs in a wireless system, the ONU-BS plays an important role in responding bandwidth requests, power allocation, and scheduling that are tackled by both the ONUs and BSs. A PON is capable of delivering bandwidth intensive voice, data, and video services at distances beyond 20 km in the network [18].

5.1.2 Integrating Mode

There have been different ways in integrating a PON and a wireless network, including peer-to-peer, overlay, and radio-over-fiber.

- **Peer-to-Peer Integration**

Peer-to-peer (P2P) integration relies on a suite of network-network interfaces (NNIs) and signaling mechanisms for the integration. It refers to a loose coupling between two networks. ONU-BSs function as border devices to interconnect a PON and a WCN [19]. A bridge function is required at the ONU-BS in order to switch between a PON frame to a WCN frame. Certainly, additional signaling is required in order to coordinate distributed ONU-BSs. Also, a longer end-to-end latency will be incurred due to bridge translation, user address mapping, and queueing delay caused by distributed design.

- **Overlay Integration**

Overlay integration is a relatively tighter coupling between the PON and WCN. In this case, each PON frame (e.g., IEEE 802.3ah frame) is segmented and encapsulated in WCN frames (e.g., IEEE 802.16m frame). The transport of WCN frames is seamless from OLT to SSs. In other words, the bridge function is not required at the ONU-

BSs, while the PON frames will be assembled and restored at each SS. The overlay integrated mode is taken as better suited to the PON-WCN due to tighter integration.

- **Radio-over-fiber Integration**

Radio-over-fiber (RoF) integration is the tightest coupling mode. The WCN signals are modulated on a wireless carrier frequency and transported in optical form between an OLT and ONU-BSs before being radiated through the air. The RoF integration also makes it easy to implement multi-cell CT among dispersed ONU-BSs.

In this study, we assume that a tight integration by way of overlay or RoF is achieved between PON and WCN. Therefore, the OLT in this case can obtain the channel status of each SS and the per-flow service dynamics. By assuming that the OLT is equipped with a strong computing unit, the integrated and centralized control over the two networks should be the trend for the next-generation hybrid access network design. Also, we envision that the PON should be designed such that the layer-2 standard can launch in line with any integrated wireless technology.

5.1.3 Cooperative Transmissions in Integrated PON-WCN

As an envisioned and desired feature of the next-generation broadband access networks, the collaborative ONU-BSs can form a virtually distributed MISO array, which allows the implementation of space-time coded transmissions to the SSs. Each cooperative ONU-BS transmits a codeword symbol which is orthogonal or quasi-orthogonal to all the others. The OLT is capable of achieving precise synchronization among the dispersed ONU-BSs, which further facilitates the CT since to keep the orthogonality of the space time codes, symbol-level synchronization of various signals from different transmitters is required at the receiver. The OLT time-multiplexes the downstream

data and broadcasts them to each dispersed ONU-BS. The ONU-BS then determines whether the data should be launched in the air under the coordination of the OLT. This is also referred to as *BU-association*, which indicates the participation of collaborative ONU-BSs for a specific user. On the other hand, a user has to be informed which frequency band(s) and codes must be used to demodulate the signal in order to successfully recover the data from the received signals. Thus, diversity gain and coding gain can be achieved due to the fact that the codeword symbols or a part of them are transmitted in parallel through different and independent wireless channels. The codewords are aggregated together and then jointly decoded at the receiver. The decoding of the codewords may take advantage of the knowledge of channels at the receiver. The channel state information can be obtained at the receiver in a way that each ONU-BSs' antenna sends training or pilot symbols to the receiver antenna so as to estimate the channel. In addition, the interplay of OFDM and virtual MISO technologies could enable interference cancelation which further improves spectra efficiency.

The multi-cell CT works in an *in-band* mode. Namely, all the ONU-BSs that participate CT to a specific SS share a common frequency band based on OFDMA technique. Note that in this research a frequency reuse factor of one is assumed in the PON-WCN deployment scenario, which is typically within a region of several kilometer diameter. For large-scale deployment in metropolitan area, a divide-and-conquer strategy can be employed, i.e., firstly, the large area is partitioned into several sections, and then in each section the DSP-PW problem is addressed respectively. The more complex issue that incorporates frequency reuse will be deferred for future research.

5.1.4 Radio Propagation and Coverage Model

The COST231-Hata model is adopted again as the radio propagation model, which is generally applicable to urban areas [83]. In this research, a site is *covered* in the area of interest (AOI) if the received radio signal strength at the site is above a given threshold level (receiver sensitivity), such that it is possible to establish a connection between the site and one or a set of ONU-BSs. The coverage is limited by link-budget, which depends on the transmission power, the receiver sensitivity, the frequency of operation, and the propagation environment. Due to the presence of buildings, trees and other obstructions, the shadowing effects still have to be considered.

With downlink multi-cell CT, two or more ONU-BSs are allowed to collaboratively provision services for an SS. The achievable rate with CT, denoted as r_{CT}, can be expressed as $r_{CT} = \frac{1}{2}\log(1 + \frac{\sum_i P_i}{N_0})$, where P_i is the power of the ith collaborative ONU-BS perceived at the SS and N_0 is the channel noise. On the other hand, in a *non-cooperative transmission* (NCT) mode, the achievable rate is denoted as r_{NCT}, and can be expressed as $r_{NCT} = \frac{1}{2}\log(1 + \frac{P}{N_0})$, where P is the power of the associated nearest ONU-BS perceived at the SS.

5.1.5 Optical Propagation Model in Fibers

Optical signals naturally lose intensity (or, attenuate) over distance. Due to the all-passiveness of PON networks, the power level of received signals at the OLT from each ONU-BS degrades with the fiber distance. The longer the distance, the lower the power level received at the OLT. Therefore, the unequal lengths of data path between the OLT and the ONU-BSs cause the so-called "*near-far*" problem, which will affect the detection of the incoming bitstream properly. If the OLT receiver can not quickly adjust its 0-1 threshold at the beginning of each received time slot, then it will

result in huge BER at the OLT receiver. Some approaches have been considered to overcome this problem [89]. Most approaches require sophisticated hardware and/or software, thereby increasing technical complexity. Alternatively, by taking "*near-far*" problem into account at the initial network planning and dimensioning stage, we may eliminate/mitigate the inherent cause of the problem at the network planning stage.

5.2 Problem Formulation

Since the integrated PON-WCN is primarily for residential and business premises, prior to the network planning, the geographical distribution of traffic demands within AOI must be known, which may be acquired by statistical data analysis of traffic measurement and/or monitoring as well as the anticipation of traffic demands growth. We define the fixed traffic-intensive sites as Very Important Points (VIPs), which may be a public hotspot, an enterprise network in an office building, an apartment block, or any densely populated place. Without loss of generality, we assume that the MSs are homogeneously distributed in the AOI, which yields uniform distributed traffic demand. The AOI is divided into a set of small rectangular grids with an equal size. A set of Test Points (TPs) specifically for the mobile coverage test are defined at each small rectangular grid. Since ONU-BSs and splitters can not be placed anywhere, only certain locations are eligible for deployment. A set of candidate positions (CPs) for allocating ONU-BSs and a set of potential sites (PSs) for allocating splitters are also defined.

Definition 5.1. Problem DSP-PW. Given the traffic demands for residential and business premises, the mobile coverage requirement in the AOI, and the cost (CAPEX and OPEX) of network equipment for the integrated PON-WCN, the design objectives of Problem DSP-PW are to:

(i) obtain the minimum overall infrastructure cost of the integrated PON-WCN;

(ii) determine the optimal locations for ONU-BSs and passive optical splitters as well as fiber layout;

(iii) decide the association between the ONU-BSs and the VIPs/TPs (i.e., *BU association*);

(iv) assign the bandwidth and power breakdown for each VIP (i.e., *RBA*).

Other inputs to the DSP-PW problem include the attenuation factor, the shadowing margins, the coverage ratio requirement, the receiver sensitivity level, the maximal transmit power of the ONU-BSs, and the upper bound of bandwidth assigned to the integrated network

Mathematical Formulation

The notations used in the problem formulation are summarized in Table 5.1. Let $\overrightarrow{\mathbb{G}} = (\mathbb{N}, \overrightarrow{\mathbb{E}})$ denote a directed graph, where $\mathbb{N}$ is the set of nodes and $\overrightarrow{\mathbb{E}}$ is the set of directed edges/links. $\mathbb{N}$ is partitioned into $\mathbb{N}_S$ and $\mathbb{N}_{CP}$, i.e., $\mathbb{N} = \mathbb{N}_S \bigcup \mathbb{N}_{CP}$, where $\mathbb{N}_S$ and $\mathbb{N}_{CP}$ represent the set of PSs for deploying splitters and the set of CPs for ONU-BSs, respectively. OLT$\in \mathbb{N}_S$. Let $\Omega = \mathbb{N}_S/\{OLT\} \bigcup \mathbb{N}_{CP}$. OLT, the splitters and the ONU-BSs are the root node, leaf nodes and internal nodes in the tree structure, respectively.

To represent such a directed tree, a binary node-edge incidence matrix $\overrightarrow{\mathbb{E}} = (e_{ij})_{|\mathbb{N}_S| \times |\Omega|}$ is defined where

$$e_{ij} = \begin{cases} 1, & \text{if there exist a directed link from } i \in \mathbb{N}_S \text{ to } j \in \Omega; \\ 0, & \text{otherwise.} \end{cases}$$

Table 5.1: Definitions of Important Symbols for Problem DSP-PW

Symbol	Definition
$\mathbb{N}_{CP}$	The set of candidate positions (CPs) for ONU-BSs within the AOI, $\lvert\mathbb{N}_{CP}\rvert = M$.
$\mathbb{N}_{TP}$	The set of test points (TPs) within the AOI, $\lvert\mathbb{N}_{TP}\rvert = N$.
$\mathbb{N}_{VIP}$	The set of very important points(VIPs) within the AOI, $\lvert\mathbb{N}_{VIP}\rvert = K$.
$\mathbb{N}_S$	The set of Potential sites (PSs) for splitters within the AOI.
$\mathbb{N}$	The set of all the nodes within AOI. $\mathbb{N} = \mathbb{N}_S \bigcup \mathbb{N}_{CP}$.
Ω	The set of the nodes within AOI except the root, $\Omega = \mathbb{N}/\{OLT\}$.
ρ_k	The traffic demand for VIP_k.
γ_0	The minimal downlink signal-to-noise-ratio (SNR) threshold for an MS.
P_{BS}	The maximal transmit power spectrum density of an ONU-BS.
B	The total available wireless bandwidth.
p_0	The transmission power spectrum density of an ONU-BS to an MS.
p_{min}	The minimal transmission power spectrum density of an ONU-BS to a VIP.
N_0	The average thermal noise power spectrum density in AOI.
η	The minimum required coverage ratio within the AOI.
c_m	The cost (CAPEX and OPEX) of an ONU-BS placed at CP_m (\$).
$\overrightarrow{\mathbb{E}}$	The set of directed edges/links in the tree, $\overrightarrow{\mathbb{E}} = (e_{ij})_{\lvert\mathbb{N}_S\rvert\times\lvert\Omega\rvert}$.
$\mathbb{D}$	The distance matrix, $\mathbb{D} = (d_{ij})_{\lvert\mathbb{N}\rvert\times\lvert\mathbb{N}\rvert}$.
L_d	The maximum allowable fiber length difference for the "near-far" problem.
$\mathbb{T}$	The set of types of splitters, $\lvert\mathbb{T}\rvert = T$.
$\mathbb{Z}$	The splitter type-location incidence matrix, $\mathbb{Z} = (z_{st})_{\lvert\mathbb{N}_s\rvert\times T}$.
$\mathbb{F}^m$	The flow-edge incidence matrix for node $m \in \mathbb{N}_{CP}$, $\mathbb{F}^m = (f_{ij}^m)_{\lvert\mathbb{N}_S\rvert\times\lvert\Omega\rvert}$.
$\overrightarrow{\mathscr{F}^m}$	The flow from OLT to node $m \in \mathbb{N}_{CP}$.
$\mathbb{R}$	The split ratio vector $\mathbb{R} = (r_t)_{1\times T}$, where r_t is the split ratio for the splitter of type t.
ψ^f	The cost per unit length of fiber (\$/m).
Ψ^s	The cost vector for splitters $\Psi^s = (\psi_t^s)_{1\times T}$, where ψ_t^s is the installation cost for the splitter of type $t \in \mathbb{T}$ (\$).
d_{ij}^f	The fiber deployment distance between node i and j.
$\mathbb{B}$	A CP location incidence vector $\mathbb{B} = (b_m)_{1\times M}$.
$\mathbb{Q}$	A TP coverage incidence vector $\mathbb{Q} = (q_n)_{1\times N}$.
$\mathbb{U}$	A VIP-CP association incidence matrix $\mathbb{U} = (\mu_{mk})_{M\times K}$.
$\mathbb{V}$	A TP-CP association incidence matrix $\mathbb{V} = (\nu_{mn})_{M\times N}$.
$\mathbb{P}$	A power breakdown assignment matrix $\mathbb{P} = (p_{mk})_{M\times K}$.
$\mathbb{W}$	A bandwidth breakdown assignment vector $\mathbb{W} = (w_n)_{1\times N}$.

A splitter type-location incidence matrix $\mathbb{Z} = (z_{st})_{|\mathbb{N}_s| \times T}$ is defined where

$$z_{st} = \begin{cases} 1, & \text{if a splitter of type } t \in T \text{ is placed at PS}_s \text{ for } s \in \mathbb{N}_S; \\ 0, & \text{otherwise.} \end{cases}$$

A directed flow incidence matrix $\mathbb{F}^m = (f_{ij}^m)_{|\mathbb{N}_S| \times |\Omega| \times |\mathbb{N}_{CP}|}$ is defined where

$$f_{ij}^m = \begin{cases} 1, & \text{if directed edge } e_{ij} \ (i \in \mathbb{N}_S, j \in \Omega) \text{ is on the path of flow } \overrightarrow{\mathscr{F}^m} (m \in \mathbb{N}_{CP}); \\ 0, & \text{otherwise.} \end{cases}$$

To formulate the DSP problem, we further define the variables as follows: A CP location incidence vector $\mathbb{B} = (b_m)_{1 \times M}$ is defined where

$$b_m = \begin{cases} 1, & \text{if CP}_m \in \mathbb{N}_{CP} \text{ is selected to place an ONU-BS;} \\ 0, & \text{otherwise.} \end{cases}$$

A TP coverage incidence vector $\mathbb{Q} = (q_n)_{1 \times N}$ is defined such that

$$q_n = \begin{cases} 1, & \text{if TP}_n \text{ is covered, i.e., } SNR_n \geq \gamma_0, \ n \in \mathbb{N}_{TP}; \\ 0, & \text{otherwise.} \end{cases}$$

$\mathbb{U} = (\mu_{mk})_{M \times K}$ represents a VIP-CP association incidence matrix where

$$\mu_{mk} = \begin{cases} 1, & \text{if the ONU-BS at CP}_m \in \mathbb{N}_{CP} \text{ provisions services to VIP}_k \in \mathbb{N}_{VIP}; \\ 0, & \text{otherwise.} \end{cases}$$

A TP-CP association incidence matrix $\mathbb{V} = (\nu_{mn})_{M \times N}$ is defined where

$$\nu_{mn} = \begin{cases} 1, & \text{if the ONU-BS at CP}_m \in \mathbb{N}_{CP} \text{ provisions services to TP}_n \in \mathbb{N}_{TP}; \\ 0, & \text{otherwise.} \end{cases}$$

We also define a transmission power breakdown assignment matrix $\mathbb{P} = (p_{mk})_{M\times K}$ and a bandwidth breakdown assignment vector $\mathbb{W} = (w_k)_{1\times K}$, where p_{mk} and w_k represent the transmission power of the ONU-BS at CP_m for VIP_k and the wireless bandwidth allocated to VIP_k, respectively. The DSP-PW problem can be formulated as follows:

$$\text{(DSP)} \qquad \underset{\mathbb{B},\mathbb{Q},\mathbb{U},\mathbb{V},\mathbb{P},\mathbb{W},\mathbb{Z},\mathbb{E},\mathbb{F}}{minimize}\ \mathcal{C} = \sum_{m=1}^{M} c_m b_m + \psi^f \sum_{i\in\mathbb{N}_S}\sum_{j\in\Omega} d_{ij}^f e_{ij} + \sum_{s\in\mathbb{N}_S}\sum_{t\in\mathbb{T}} \psi_t^s z_{st} \tag{5.1}$$

subject to the following constraints:

$$w_k \log(1 + \frac{1}{N_0}\sum_{m=1}^{M}\frac{p_{mk}}{d_{mk}^{\alpha}}) \geq \rho_k, \quad \forall k \in \mathbb{N}_{VIP} \tag{5.2}$$

$$\frac{1}{N}\sum_{n=1}^{N} q_n \geq \eta \times 100\% \tag{5.3}$$

$$\sum_{m=1}^{M} \mu_{mk} \geq 1, \quad \forall k \in \mathbb{N}_{VIP} \tag{5.4}$$

$$\mu_{mk} \leq b_m, \quad \forall m \in \mathbb{N}_{CP}, \forall k \in \mathbb{N}_{VIP} \tag{5.5}$$

$$\sum_{m=1}^{M} \nu_{mn} \geq 1, \quad \forall n \in \mathbb{N}_{TP} \tag{5.6}$$

$$\nu_{mn} \leq b_m, \quad \forall m \in \mathbb{N}_{CP}, \forall n \in \mathbb{N}_{TP} \tag{5.7}$$

$$q_n \geq \frac{1}{\sum_{m=1}^{M}\frac{p_0}{N_0 d_{mn}^{\alpha}} - \gamma_0}(\sum_{m=1}^{M}\frac{v_{mn}p_0}{N_0 d_{mn}^{\alpha}} - \gamma_0), \forall n \in \mathbb{N}_{TP} \tag{5.8}$$

$$1 - q_n \geq \frac{1}{\sum_{m=1}^{M}\frac{p_0}{N_0 d_{mn}^{\alpha}} - \gamma_0}(\gamma_0 - \sum_{m=1}^{M}\frac{v_{mn}p_0}{N_0 d_{mn}^{\alpha}}), \forall n \in \mathbb{N}_{TP} \tag{5.9}$$

$$\sum_{k=1}^{K} p_{mk} < P_{BS}, \quad \forall m \in \mathbb{N}_{CP} \tag{5.10}$$

$$\mu_{mk} \geq p_{mk}/P_{BS}, \quad \forall m \in \mathbb{N}_{CP}, \forall k \in \mathbb{N}_{VIP} \tag{5.11}$$

$$p_{mk} \geq \mu_{mk} p_{min}, \quad \forall m \in \mathbb{N}_{CP}, \forall k \in \mathbb{N}_{VIP} \tag{5.12}$$

$$\sum_{k=1}^{K} w_k < B, \tag{5.13}$$

$$b_m, q_n, \mu_{mk}, \nu_{mn} \in \{0, 1\}, \quad \forall m \in \mathbb{N}_{CP}, n \in \mathbb{N}_{TP}, k \in \mathbb{N}_{VIP} \tag{5.14}$$

$$p_{mk} \geq 0, \quad \forall m \in \mathbb{N}_{CP}, \quad \forall k \in \mathbb{N}_{VIP} \tag{5.15}$$

$$w_k > 0, \quad \forall k \in \mathbb{N}_{VIP} \tag{5.16}$$

The objective function (5.1) minimizes the total cost. Constraint (5.2) ensures that the traffic demand for each VIP is satisfied; i.e., the throughput for each VIP is no smaller than its traffic load demand. Constraint (5.3) ensures the coverage ratio requirement. Constraints (5.4) and (5.6) are the BU association constraints, by which the VIPs/TPs must be associated with at least one ONU-BS since more than one collaborative ONU-BSs may provide more reliable high data rate transmissions. Constraints (5.5) and (5.7) ensure that an ONU-BS must place at the CP_m if it is associated with a VIP/TP. Constraints (5.8) and (5.9) stipulate the definition of $\mathbb{Q}$. Constraints (5.10) and (5.13) define the upper bound for an ONU-BS's power and bandwidth, respectively. Constraints (5.11) and (5.12) stipulate that if $p_{mk} = 0$ then $\mu_{mk} = 0$, and vice versa. Constraints (5.14)-(5.16) state that each entry in $\mathbb{B}$, $\mathbb{Q}$, $\mathbb{U}$ and $\mathbb{V}$ is binary, and that of $\mathbb{P}$ and $\mathbb{W}$ is non-negative and positive, respectively.

To construct a directed tree for the integrated network, the inherent topological features of (i) *acyclic*, (ii) *connected* and (iii) *directed* must be maintained. To capture those features, we formulate the node-edge incidence matrix and splitter type-location

incidence matrix subject to the following constraints:

$$e_{ij} + e_{ji} \leq 1 \qquad \forall i, j \in \mathbb{N}_S / \{OLT\} \tag{5.17}$$

$$\sum_{i \in \mathbb{N}_S} e_{ij} = \sum_{t \in \mathbb{T}} z_{jt} \qquad \forall j \in \mathbb{N}_S / \{OLT\} \tag{5.18}$$

$$\sum_{i \in \mathbb{N}_S} e_{ij} = b_j \qquad \forall j \in \mathbb{N}_{CP} \tag{5.19}$$

$$\sum_{j \in \Omega} e_{ij} \leq \sum_{t \in \mathbb{T}} z_{it} r_t \qquad \forall i \in \mathbb{N}_S \tag{5.20}$$

$$\sum_{j \in \Omega} e_{ij} \geq \sum_{t \in \mathbb{T}} z_{it} \qquad \forall i \in \mathbb{N}_S / \{OLT\} \tag{5.21}$$

$$e_{ij} \in \{0, 1\} \qquad \forall i \in \mathbb{N}_S, j \in \Omega \tag{5.22}$$

$$\sum_{t \in \mathbb{T}} z_{it} \leq 1 \qquad \forall i \in \mathbb{N}_S \tag{5.23}$$

$$z_{st} \in \{0, 1\} \qquad \forall s \in \mathbb{N}_S, t \in \mathbb{T} \tag{5.24}$$

Constraint (5.17) stipulates that the tree is a directed one, namely, if $e_{ij} = 1$ node i is the parent of node j. Constraints (5.18) ensures each internal node (i.e., PS) has exactly one parent if it is in the directed tree and (5.19) ensures each leaf node (i.e., ONU-BS) also has exactly one parent. Constraint (5.20) ensures the number of child nodes for each splitter is upper bounded by its split ratio. Constraint (5.21) ensures that the internal node has at least one outgoing edges if it is in the tree, without this

constraint, the splitter will be a leaf node.

$$f_{ij}^{m} \leq e_{ij} \qquad \text{for } \forall i \in \mathbb{N}_S, j \in \Omega, m \in \mathbb{N}_{CP} \tag{5.25}$$

$$\sum_{i \in \Omega} f_{OLT,i}^{m} = b_m \qquad \text{for } \forall m \in \mathbb{N}_{CP} \tag{5.26}$$

$$\sum_{i \in \mathbb{N}_S} f_{i,m}^{m} = b_m \qquad \text{for } \forall m \in \mathbb{N}_{CP} \tag{5.27}$$

$$\sum_{j \in \mathbb{N}_S, j \neq i} f_{ji}^{m} = \sum_{v \in \Omega, v \neq i} f_{iv}^{m} \quad \text{for } \forall i \in \mathbb{N}_S / \{OLT\}, m \in \mathbb{N}_{CP} \tag{5.28}$$

$$-L_d \leq \sum_{i \in \mathbb{N}_S} \sum_{j \in \Omega} d_{ij} f_{ij}^{m_1} - \sum_{i \in \mathbb{N}_S} \sum_{j \in \Omega} d_{ij} f_{ij}^{m_2} \leq L_d \tag{5.29}$$

$$\text{for } \forall m_1, m_2 \in \mathbb{N}_{CP}, m_1 \neq m_2$$

$$f_{ij}^{m} \in \{0, 1\} \qquad \text{for } \forall i \in \mathbb{N}_S, j \in \Omega, m \in \mathbb{N}_{CP} \tag{5.30}$$

To formulate the "near-far" problem constraint, we employ a logical *flow* concept. More specifically, for a leaf node, say $m \in \mathbb{N}_{CP}$, there exists one corresponding flow, denoted as $\overrightarrow{\mathscr{F}^m}$. The source of the flow $\overrightarrow{\mathscr{F}^m}$ is the root (OLT) which only generates one flow, but does not receive flow. The leaf node m only receives the flow, but does not generate or relay any flow. The internal node except the root (OLT) and node m obey flow conservation if they are along the path of flow $\overrightarrow{\mathscr{F}^m}$, i.e., the number of incoming flows equals that of outgoing flows.

Constraint (5.25) ensures that the flow $\overrightarrow{\mathscr{F}^m}$ at e_{ij} exists only when there exists an edge first, which also indicates that the flow is a directed one. Constraints (5.26) and (5.27) ensure that the root and leaf have only one unit flow incidence. Constraint (5.28) formulates the flow conservation property at internal node m. The "near-far" constraint is formulated in ((5.29)), which is optional for the DSP problem. It ensures that the flow length difference between any two flows is less than a predefined value L_d. If the network designer prefers alternative approach to solve the "near-far" problem,

(5.29) can be removed.

The above constraints ensure a directed tree as explained below. (i) Regarding the *acyclic* property, the definition of $\mathbb{E}$ stipulates there exist no incoming edges towards the root node (i.e., $e_{i,OLT}$ for $i \in \mathbb{N}$), or outgoing edges away from leaf nodes (i.e., e_{mi} for $m \in \mathbb{N}_{CP}$, $i \in \mathbb{N}_S$). Constraints (5.19, 5.22, 5.25, 5.26, 5.27) further ensure that the degree of leaf nodes equals one, thus there exist no loops in the graph. (ii) Regarding the *connected* property, there exist exactly one incoming edge and at least one outgoing edges from the internal nodes (constraints 5.18, 5.21, 5.22, 5.25, 5.28, 5.30). (iii) Constraints (5.17) and (5.25) guarantee that both the edges and flows in the tree are *directed*.

5.3 Solution

5.3.1 Problem Decomposition

Due to the large amount of decision variables, finding an optimal solution of Problem DSP-PW is challenging since the search space is exponentially increased with the number of decision variables. To reduce the computation complexity, we decompose the DSP-PW problem into two subproblems that are sequentially solved as shown in Figure 5.2, i.e., Subproblem 1 (DSP-W) and Subproblem 2 (DSP-P), respectively. Specifically, Subproblem 1 is to minimize the total infrastructure cost for ONU-BS deployment (i.e., the 1st term in Eq. 5.1) while both the traffic demand for each VIP and the coverage requirement in the AOI can be satisfied. Constraints (5.2)-(5.16) form the DSP-W formulation. Based on the resulting ONU-BS positions in Subproblem 1, Subproblem 2 is to minimize the total cost for PON deployment (i.e., the 2nd and 3rd terms in Eq. 5.1), in which the outputs will be the fiber layout and the splitter placement. Constraints

(5.17)-(5.30) form the DSP-P formulation. Thus, after the decomposition, the search spaces of Subproblems 1 and 2 can be significantly reduced.

The Subproblem 2 (DSP-P) is an MILP, while Subproblem 1 (DSP-W) is an MINLP which can hardly be solved by any systematic approach [11]. In the next subsection, we propose an approach to reformulate the Subproblem 2 (DSP-P) into a solvable MILP.

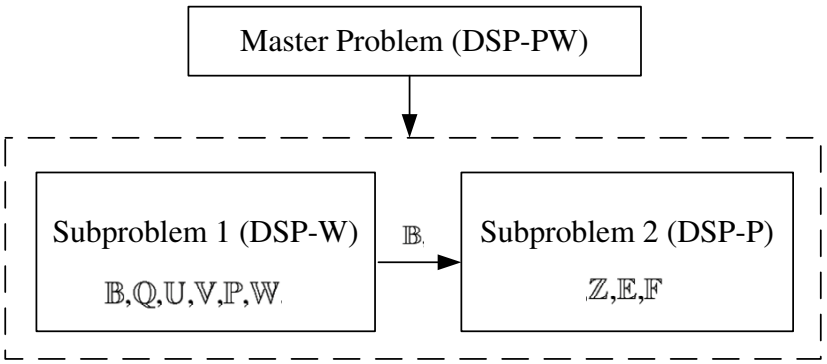


Figure 5.2: Model decomposition structure of problem (DSP-PW).

5.3.2 Linear Approximation Based Reformulation

We observe that only constraint (5.2) is nonlinear in the DSP-W formulation. To reformulate it as an MILP, we employ a linear approximation approach. And the approximation error is manipulated to be upper bounded by ε so that we can achieve a precise linear approximation. The approach is elaborated as follows.

We define a new working variable ξ_k $(\forall k = 1 \dots K)$ such that

$$\xi_k = \sum_{m=1}^{M} \frac{1}{N_0 \cdot d_{mk}^{\alpha}} \cdot p_{mk}$$

.

In practise the value of ξ_k is in $[\xi_{min}^k, \xi_{max}^k]$, where $\xi_{min}^k = \frac{2P_{min}}{N_0 \max_{\forall m} d_{mk}^{\alpha}}$, and $\xi_{max}^k = \frac{MP_{BS}}{N_0 \min_{\forall m} d_{mk}^{\alpha}}$. With variable ξ_k, constraint (5.2) can be formulated as

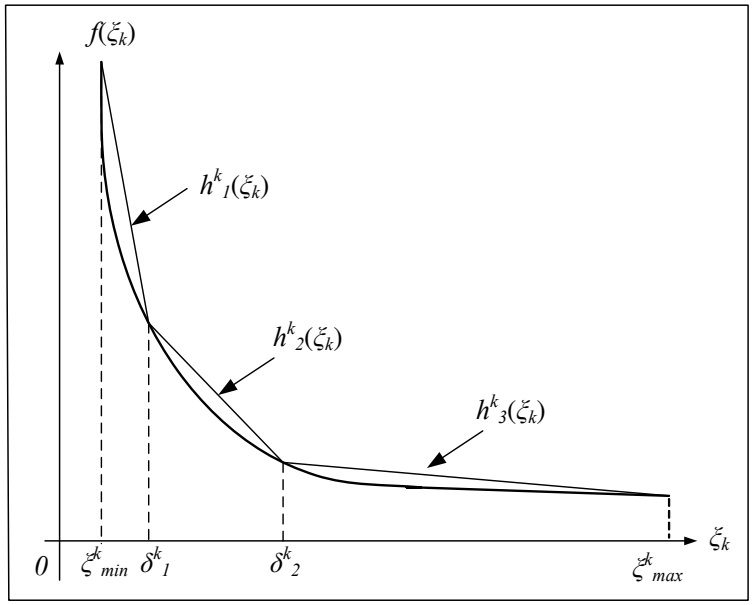


Figure 5.3: Linear approximation of the convex function $f(\xi_k)$

$$w_k \log(1+\xi_k) \geq \rho_k, \qquad \text{for } \forall k = 1 \ldots K$$

which can be reordered as

$$w_k \geq \frac{\rho_k}{\log(1+\xi_k)}, \qquad \text{for } \forall k = 1 \ldots K \tag{5.31}$$

Let $f(\xi_k) = \frac{\rho_k}{\log(1+\xi_k)}$, which is a convex function. We divide the interval of ξ_k into $l+1$ regions, i. e., $[\xi^k_{min}, \xi^k_{max}] \rightarrow [\xi^k_{min}, \delta^k_1], [\delta^k_1, \delta^k_2], \ldots, [\delta^k_{i-1}, \delta^k_i], \ldots, [\delta^k_l, \xi^k_{max}]$, where δ^k_i represents the end of region i. Note that the size of each region could be different. Let δ^k_i and δ^k_{i+1} be the values of ξ_k where the chord intersects the nonlinear convex function ($\delta^k_i < \delta^k_{i+1}$). The steepness (or, slope) of the line segment $h^k_i(\xi_k)$ between points $(\delta^k_{i-1}, f(\delta^k_{i-1}))$ and $(\delta^k_i, f(\delta^k_i))$ can be expressed as

$$h^k_i = \frac{f(\delta^k_{i+1}) - f(\delta^k_i)}{\delta^k_{i+1} - \delta^k_i} = \left(\frac{1}{\log(1+\delta^k_i)} - \frac{1}{\log(1+\delta^k_{i+1})} \right) \frac{\rho_k}{\delta^k_{i+1} - \delta^k_i} \tag{5.32}$$

Therefore, the linear approximation of function $f(\xi_k)$ is expressed as follows:

$$f(\xi_k) = \frac{\rho_k}{\log(1+\xi_k)} \tag{5.33}$$

$$\approx \begin{cases} h_1^k(\xi_k) = a_1^k(\xi_k - \xi_{min}^k) + f(\xi_{min}^k), & \text{for } \xi_{min}^k \le \xi_k \le \delta_1^k; \\ h_2^k(\xi_k) = a_2^k(\xi_k - \delta_1^k) + f(\delta_1^k), & \text{for } \delta_1^k \le \xi_k \le \delta_2^k; \\ \quad \ldots\ldots & \\ h_l^k(\xi_k) = a_l^k(\xi_k - \delta_l^k) + f(\delta_{l-1}^k), & \text{for } \delta_{l-1}^k \le \xi_k \le \delta_l^k; \\ h_{l+1}^k(\xi_k) = a_{l+1}^k(\xi_k - \xi_{max}^k) + f(\delta_l^k), & \text{for } \delta_l^k \le \xi_k \le \xi_{max}^k. \end{cases} \tag{5.34}$$

The values of δ_i^ks have a critical impact on the linear approximation of $f(\xi_k)$. In order to obtain an accurate linear approximation of the curve $f(\xi_k)$, we provide an algorithm to determine δ_i^ks such that the approximation error (i.e., the chord-curve distance) is manipulated to be less than ε. The algorithm is shown in Figure 5.4.

The inputs of the algorithm include ε, ξ_{min}^k, ξ_{max}^k and function $f(\xi_k)$. The output is *Splitpoints*, the set of all δ_i^ks, by which we can obtain all the line segments to approximate the curve $f(\xi_k)$ with Eq. (5.34). The function *LineCurveDistance*(δ_1^k, δ_2^k) in the algorithm is to obtain the distance, denoted by D, between the curve $f(\xi_k)$ and the line segment defined by points $(\delta_1^k, f(\delta_1^k))$ and $(\delta_2^k, f(\delta_2^k))$. Note that $\varepsilon \in (0, D_{UB}]$, where D_{UB} is the maximal distance between the curve and the line defined by points $(\xi_{min}^k, f(\xi_{min}^k))$ and $(\xi_{max}^k, f(\xi_{max}^k))$. The calculation of D_{UB} is given in Appendix I, and *LineCurveDistance* can also follow similar steps to yield D. If $D > \varepsilon$, then we set δ_2^k as $(\delta_1^k + \delta_2^k)/2$; otherwise, δ_2^k and δ_1^k are set as an entry of *Splitpoints* and δ_2^k, respectively. The loop will terminate when $\delta_1^k = \delta_2^k$. The algorithm can effectively manipulate the linear approximation error upper bounded by ε.

- **MILP formulation**

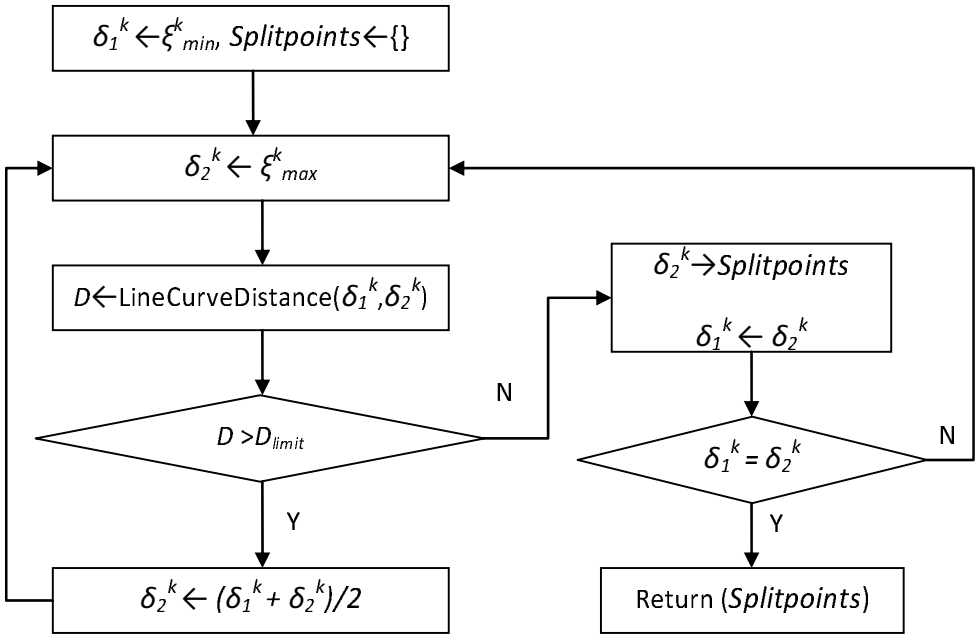


Figure 5.4: Flow chart of the ε-approximation algorithm to linearize curve $f(\xi_k)$.

With the above linearization process, the nonlinear constraint (5.2) in the original formulation can be replaced with the following constraints.

$$\xi_k = \sum_{m=1}^{M} \frac{p_{mk}}{N_0 d_{mk}^{\alpha}} \quad \text{for } \forall k = 1 \dots K \tag{5.35}$$

$$\xi_{min}^k \leq \xi_k \leq \xi_{max}^k \quad \text{for } \forall k = 1 \dots K \tag{5.36}$$

$$\omega_k \geq \frac{f(\delta_1^k) - f(\xi_{min}^k)}{\delta_1^k - \xi_{min}^k}(\xi_k - \xi_{min}^k) + f(\xi_{min}^k) \text{for } \xi_k \in [\xi_{min}^k, \delta_1^k], \forall k = 1 \dots K, \tag{5.37}$$

$$\omega_k \geq \frac{f(\delta_2^k) - f(\delta_1^k)}{\delta_2^k - \delta_1^k}(\xi_k - \delta_1^k) + f(\delta_1^k) \quad \text{for } \xi_k \in [\delta_1^k, \delta_2^k], \forall k = 1 \dots K, \tag{5.38}$$

$$\dots\dots \tag{5.39}$$

$$\omega_k \geq \frac{f(\delta_l^k) - f(\delta_{l-1}^k)}{\delta_l^k - \delta_{l-1}^k}(\xi^k - \delta_{l-1}^k) + f(\delta_{l-1}^k) \text{for } \xi_k \in [\delta_{l-1}^k, \delta_l^k], \forall k = 1 \dots K, \tag{5.40}$$

$$\omega_k \geq \frac{f(\xi^k_{max}) - f(\delta^k_l)}{\xi^k_{max} - \delta^k_l}(\xi^k - \delta^k_l) + f(\delta^k_l) \quad \text{for } \xi_k \in [\delta^k_l, \xi^k_{max}], \forall k = 1 \ldots K. \tag{5.41}$$

Constraints (5.35)-(5.41) and (5.3)-(5.16) form an MILP which can nicely approximate the original problem and can be solved by an MILP solver.

5.4 Numerical Results and Case Studies

In this section, case studies are conducted to validate the proposed problem formulation and optimization method. We also demonstrate the performance benefits by adopting cooperative service provisioning in integrated PON-WCN. Before going into more practical scenarios, the performance gain in a simple case with three cooperatively provisioning ONU-BSs is investigated, aiming to provide an insight on the fundamental performance behavior of multi-cell CT. Figure 5.5 compares the achievable rate between CT and NCT modes. The coordinates in x-axis and y-axis represent normalized distance. The number shown on each contour line represents the corresponding value of achievable rate when the receiver moves along the contour line. It is observed that the CT mode outperforms the NCT mode in terms of the improvement of achievable rate or coverage extension, especially at the cell boundaries.

To simulate practical deployment, three scenarios are investigated, where an OFDMA-based IEEE 802.16m air interface is assumed in the experiments. We assume all the CPs with the same CAPEX and OPEX for placing an ONU-BS. The maximal transmission power of the ONU-BSs and the channel attenuation factor are set as 10w and 3, respectively. The radio bandwidth allocated to the system is 20MHz. In addition, the external interference is assumed to be constant and embedded into thermal noises. Again CPLEX 11.0 [84] is employed for solving the reformulated DSP-PW problem. We use a workstation with an Intel Zeon 3GH CPU and 4GB RAM. Figure 5.6 illus-

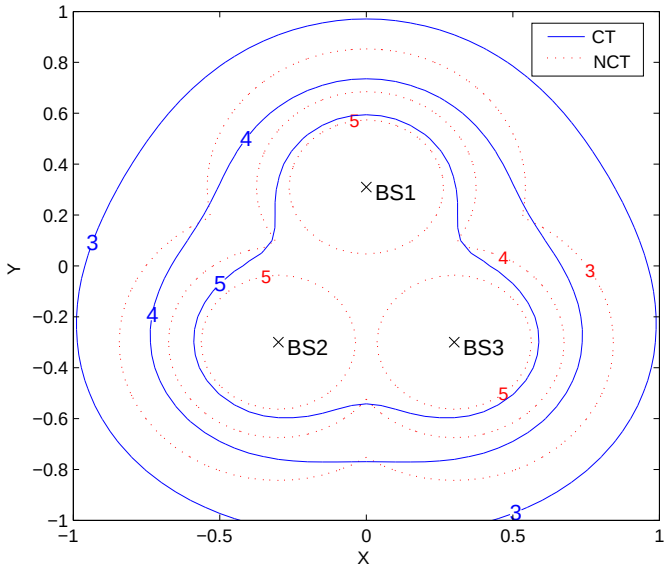


Figure 5.5: 2-D contour of achievable rate: comparison in CT and NCT mode.

trates the network layout of scenario (I). The coordinate of each VIP in the AOI is normalized, and the amount of traffic demand is proportional to the radius of the circle representing the VIP.

The problem size, the number of constraints, variables, average computation time and average optimality gap in the simulated scenarios are shown in Table 5.2, respectively. With the simulation, we hope to verify the formulation and the solvability of the DSP problem. Table 5.2 indicates that the problem size grows dramatically as the network size increases, and the problem can be solved successfully with the decomposition approach. Without decomposition, the computation time is intolerable for solving the DSP problem, especially for larger size network.

Note that Scenario (III) reflects a real-world PON-WCN deployment scenario of several kilometers diameter area where the distance between two neighboring TPs is

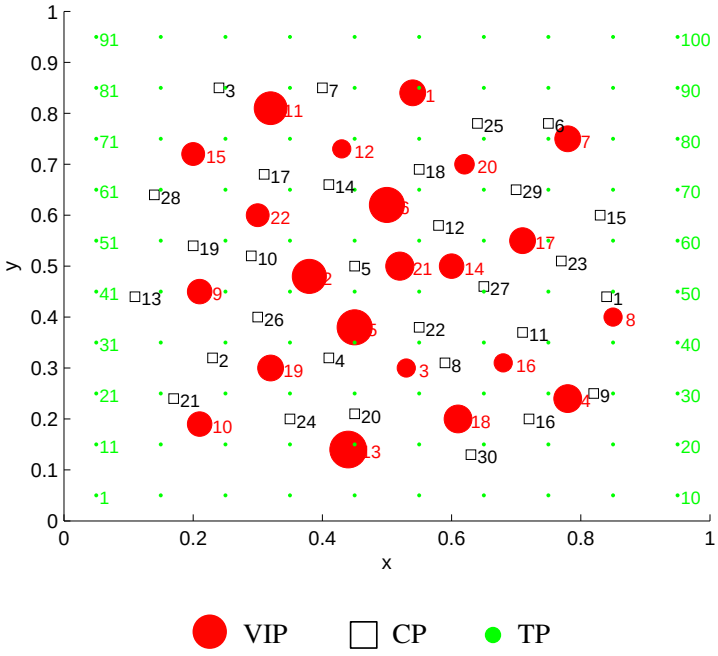


Figure 5.6: Illustration of network layout of AOI in Scenario (I).

around several hundred meters. The results of Scenario (III) demonstrate the effectiveness and efficiency of the proposed formulation and optimization method in practical network planning tasks.

For the Subproblem (DSP-W), we are interested in observing the results of BU-association in terms of the number of associated ONU-BSs for each VIP and the mobile coverage quality (i.e., the perceived SNR at each TP) with respect to different SNR thresholds. Moreover, the scalability in capacity enhancement is of particular interest, namely, the change of objective value when the traffic demands of the VIPs are increased. For these purposes, we define an increasing ratio of total traffic demands for all the VIPs, denoted as $G_\rho = \frac{\sum_{k=1}^{K} \rho'_k}{\sum_{k=1}^{K} \rho_k}$, where ρ_k and ρ'_k are the original and the new traffic demand of VIP$_k$, respectively.

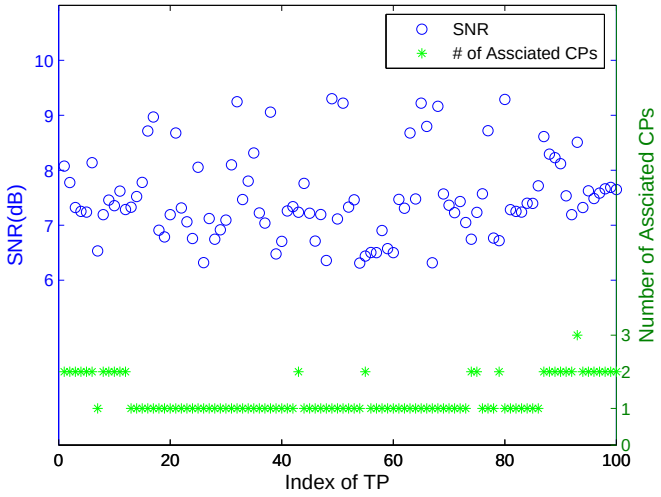
Table 5.2: Problem Size, Average Computation Time and Average Optimality Gap for CPLEX Solving Problem DSP-PW

Scenario	Number of Nodes				Constraints	Variables
	VIP	CP	TP	PS		
(I)	22	30	100	35	39435	42847
(II)	40	60	225	75	327790	333559
(III)	80	100	625	105	1629593	1868450

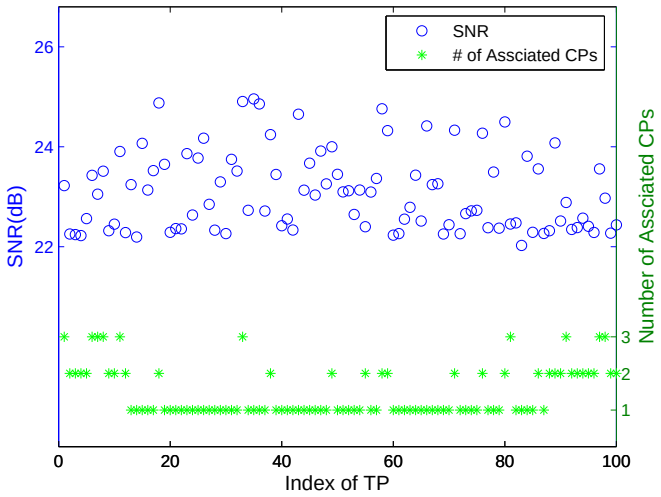
Scenario	Ave. computation time		Ave. optimality gap	
	(DSP-W)	(DSP-P)	(DSP-W)	(DSP-P)
(I)	48.6 Sec	106.2 Sec	2.18%	0.00%
(II)	73.4 Sec	1469.2 Sec	3.96%	1.03%
(III)	801.4 Sec	21594 Sec	3.14 %	0.11%

Table 5.3 demonstrates the result of BU-association and spectra efficiency for each VIP by taking the CT mode and NCT mode in Scenario (I), respectively. With CT, the spectra efficiency gains, denoted as $G_{rate} = \frac{r_{CT}-r_{NCT}}{r_{NCT}}$, for the VIPs range from 8.99% to 85.71%, which then contributes to the overall system throughput enhancement due to the participation of more BSs in the cooperative transmissions. As a result, to satisfy the same traffic demands, 13 BSs are required to be deployed at CP_4-CP_7, CP_9, CP_{10}, CP_{18}-CP_{21}, CP_{23}, CP_{27} and CP_{30} in the CT mode; while in NCT mode, 16 BSs are required to be placed at CP_1, CP_3-CP_5, CP_6, CP_8, CP_9, CP_{12},CP_{14}, CP_{18}, CP_{20},CP_{21}, CP_{23}, CP_{26}, CP_{27} and CP_{30}, respectively. The total deployment cost for ONU-BSs is reduced by 18.75% in the CT mode compared with the NCT mode.

Figures 5.7 (a) and (b) show the results in terms of user-perceived SNR and BU-



(a) $\gamma_0 = 6$dB



(b) $\gamma_0 = 22$dB

Figure 5.7: The perceived SNR and the number of associated CPs for each TP in Scenario (I).

association at each TP in Scenario (I) with the SNR threshold as 6dB and 22dB, respectively [90]. The TPs' perceived SNR values reflect the mobile coverage quality within the AOI. We set the coverage ratio to 95%. It is observed that a majority of TPs only need to be associated with one BS for both Figure 5.7 (a) and (b). On the other hand, for the TPs at the edge of the AOI such as TP_1-TP_{11} and TP_{90}-TP_{100}, two or three ONU-BSs are associated with a TP to meet the SNR threshold by way of CT. Moreover, by comparing 5.7 (a) and (b), we further observe that the number of TPs that are associated with 2 or 3 ONU-BSs in Figure 5.7(a) is much more than that in Figure 5.7(b). It indicates that the improvement of mobile coverage quality is gained at the expenses of the participation of more ONU-BSs collaboratively transmit to an SS (TP).

Figure 5.8 shows the key results of the case studies, i.e., the change of objective value (the number of ONU-BSs) with the increase of traffic demands of VIPs in Scenario (I), (II) and (III). It is observed that when the traffic demands are small, both CT and NCT mode yield the same objective value. The reason is that the traditional NCT mode can already achieve the design requirements. However, with the growth of traffic demands, e.g. $G_\rho \in [1.6, 3.8]$ in Scenario (I), the number of ONU-BSs in CT mode is always less than that in NCT mode (9.09% to 27.27%), which demonstrates the advantages of CT in terms of significant cost saving. Moreover, by adopting CT, the system can satisfy even higher traffic demands than the NCT mode. Specifically, as shown in Figure 5.8(a) of Scenario (I), when G_ρ goes up to over 3.8, the system can not achieve the requirement in NCT mode; however, it can be achieved by placing 16 ONU-BSs in CT mode. Moreover, the system can still meet the traffic demand growth when G_ρ reaches 11.2 by placing 30 ONU-BSs. Similarly, in Scenario (II) and (III), CT mode also outperforms NCT mode significantly. Therefore, Figure 5.8 not only

demonstrates a salient advantageous economic scale of CT, but also the scalability in capacity enhancement for the ever-increasing future residential and business premises, which on the other hand verifies the cost-effectiveness of the proposed PON-WCN architecture.

For the Subproblem 2 (DSP-P), we focus on validating the formulation. For this purpose, a relative small size network is simulated where the locations of ONU-BSs are obtained by solving Subproblem 1 (DSP-W) in Scenario (I), as shown in Figure 5.9. L_d is set to 2km in our simulations [89]. We use the manhattan distance between each pair of nodes to represent the link cost. Figure 5.10 shows the result of fiber deployment and splitter placement. It is observed that the tree structure is maintained with the topological features of acyclic, connected and directed, where the OLT, splitters, and ONU-BSs function as the root, internal and leaf nodes, respectively. The directed flow towards each leaf node is listed as follows:

$\overrightarrow{\mathscr{F}^1}$: OLT$\rightarrow$ ONU-BS$_1$. $\overrightarrow{\mathscr{F}^2}$: OLT$\rightarrow$ ONU-BS$_2$.

$\overrightarrow{\mathscr{F}^3}$: OLT$\rightarrow$ PS$_1$ $\rightarrow$ PS$_2$ $\rightarrow$ ONU-BS$_3$.

$\overrightarrow{\mathscr{F}^4}$: OLT$\rightarrow$ PS$_1$ $\rightarrow$ PS$_2$ $\rightarrow$ ONU-BS$_4$.

$\overrightarrow{\mathscr{F}^5}$: OLT$\rightarrow$ PS$_3$ $\rightarrow$ ONU-BS$_5$.

The resultant maximal distance difference among $\overrightarrow{\mathscr{F}^m}$s ($m = 1 \ldots 5$) is 1.82km, which is less than L_d. Hence, we have verified that the the formulation of (DSP-P) not only can ensure a tree topology, but also can solve the "near-far" problem.

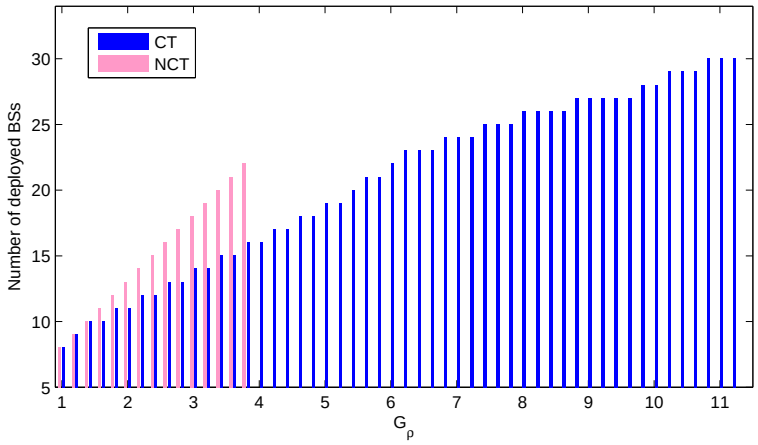
Obviously, under our proposed network architecture, as the cost for ONU-BSs has been reduced in subproblem 1 (DSP-W) at the first stage, the resultant overall infrastructure cost can be reduced significantly over conventional network with NCT mode while meeting the same long-term performance requirements.

5.5 Summary

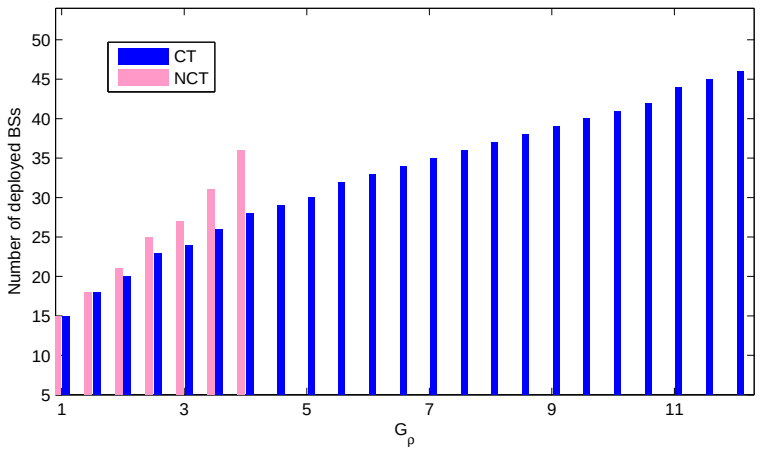
In this chapter, we have studied the problem of network dimensioning and site planning in a novel integrated PON-WCN architecture for FMC. The integrated architecture is identified to have economic and performance merits by taking advantage of multi-cell CT. The problem leads to a novel optimization framework. We have provided a solution based on decomposition and linear approximation to solve the complex problem effectively. Case studies have shown a significant infrastructure cost reduction, spectra efficiency improvement and scalability in capacity enhancement under multi-cell CT. Given a set of network parameters, the proposed optimization framework can provide design guidelines on the network deployment and cost estimations.

Table 5.3: BU-Association and Spectra Efficiency of VIPs in Scenario (I) for CPLEX Solving Problem DSP-PW

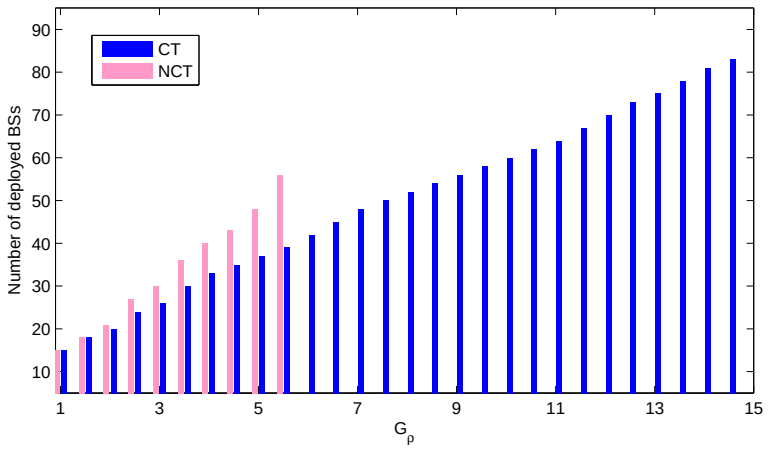
VIP	Index of Associated CP		Spectra Efficiency		
	CT	NCT	CT	NCT	G_{rate}
1	CP_6, CP_9	CP_6	3.762	2.072	81.64%
2	CP_4, CP_5, CP_{10}	CP_5	4.243	2.634	61.22%
3	CP_4, CP_{20}	CP_8	4.376	3.985	9.80%
4	CP_9,CP_{23}	CP_9	4.315	2.592	66.41%
5	CP_4, CP_5, CP_{20}	CP_4	4.496	3.665	22.68%
6	CP_5, CP_7, CP_{18}	CP_{12}	4.216	3.572	17.93%
7	CP_6, CP_{23}	CP_6	4.138	2.903	42.41%
8	CP_9, CP_{23}	CP_1	4.263	4.082	4.41%
9	CP_{10}, CP_{19}	CP_{26}	4.315	2.637	63.88%
10	CP_{20}, CP_{21}	CP_{21}	4.153	3.982	4.27%
11	CP_7, CP_8, CP_{19}	CP_3	4.556	2.607	74.33%
12	CP_7, CP_{18}	CP_{14}	4.415	2.858	54.20%
13	CP_4, CP_{20}, CP_{21}	CP_{20}	4.358	3.624	20.17%
14	CP_{23}, CP_{27}	CP_{27}	3.891	3.419	14.08%
15	CP_{10}, CP_{19}	CP_3	3.777	2.654	42.26%
16	CP_9, CP_{30}	CP_9	3.950	2.766	42.60%
17	CP_6, CP_{23}	CP_{23}	3.882	3.564	8.99%
18	CP_9, CP_{30}	CP_{30}	3.919	3.423	14.62%
19	CP_4, CP_{21}	CP_{26}	4.419	2.617	68.32%
20	CP_6, CP_7, CP_{18}	CP_{18}	4.054	3.323	21.99%
21	CP_5, CP_{27}	CP_5	4.135	2.393	72.80%
22	CP_{10}, CP_{19}	CP_{14}	4.027	2.176	85.71%



(a) Scenario (I)



(b) Scenario (II)



(c) Scenario (III)

Figure 5.8: The number of ONU-BSs vs. the increasing ratio of total traffic demands of all VIPs.

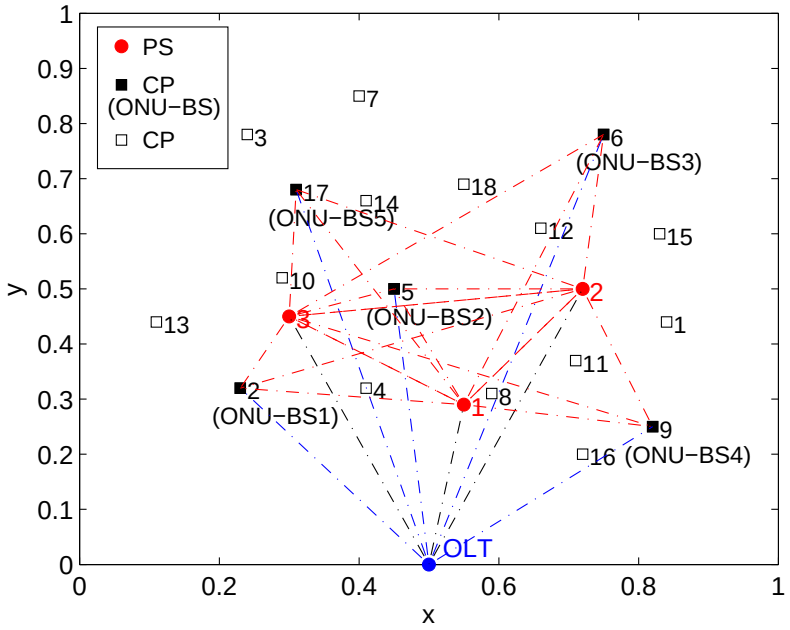


Figure 5.9: The layout of PSs, CPs, and obtained ONU-BSs after solving Subproblem 1 (DSP-W) in Scenario (I).

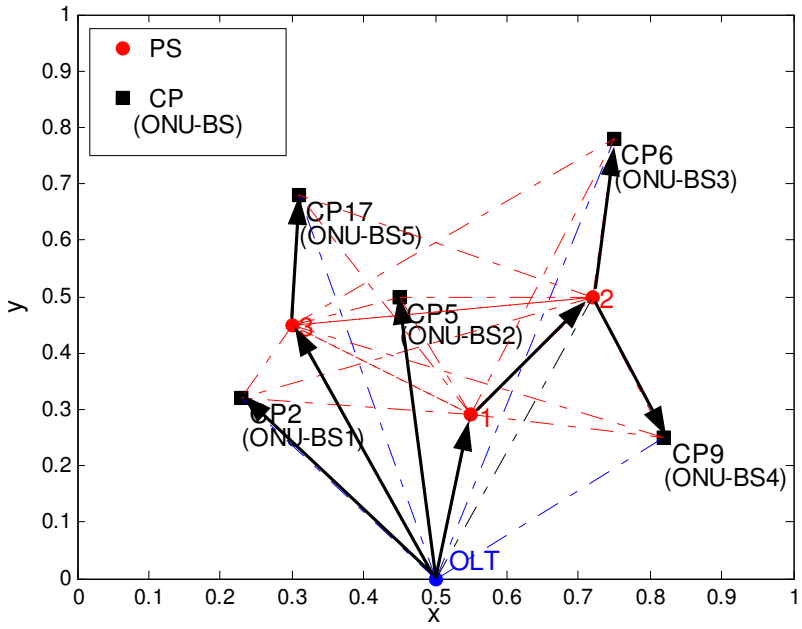


Figure 5.10: The resultant optimal PON tree (including fiber layout and splitter placement) in Scenario (I).

Chapter 6

Conclusions and Future Research

In this chapter, the contributions of this dissertation are concluded, followed by the future research directions and final remarks.

6.1 Major Research Contributions

This dissertation focuses on the issue of location planning in broadband access networks under cooperative transmission. The insight of this research comes from the following facts: one is the salient benefits from the advances in wireless cooperative communications, the other is the strategic impacts of location decisions. The recognition of the significance by incorporating both to achieve a cost-effective design has led the direction of this research, which has been seldom explored previously.

We study the placement problem in two categories of network deployment environment: an existing wireless access network and a perspective broadband access network.

For the first deployment environment, to solve some practical problems such as the requirements of capacity enhancement and coverage extension in an existing cell, RS has been introduced in the network architecture. Based on the single-cell cooperative

relay technology, we have proposed two optimization frameworks with the design objectives of maximizing cell capacity and minimizing number of RSs for deployment, respectively. Mathematical formulations have been provided to precisely capture the characteristics of the placement problems as well as the behavior under multi-level cooperative relaying. The corresponding solution algorithms have been developed to obtain the optimal (or near-optimal) results in polynomial time. Numerical analysis and case studies have been conducted to validate the performance benefits due to RS placement.

For the second environment, to deploy a new metropolitan broadband access network, we focus on integrating the state-of-the-art optical and wireless networks as well as employing multi-cell cooperative transmission technology. We have provided an optimization framework to solve the problem of dimensioning and site planning in an integrated PON-WCN. Besides node placement, BS-user association, wireless bandwidth and power breakdown assignment have been jointly considered in a single stage to achieve better performance. We have also proposed a solution to the complex optimization problem based on decomposition and linear approximation. Numerical analysis and case studies have been conducted to verify the the proposed framework and performance and economic benefits.

6.2 Future Research

The research in this dissertation focuses on the location planning in broadband access networks for the residential and business premises of wireless and mobile services. In this section, we identify several research directions for future work as follows.

- Joint Design of Location Planning and Frequency Planning.

For a large-scale wireless access network which covers metropolitan areas, an opti-

mal *frequency reuse* scheme is important to increase the capacity of a wireless system. Although frequency planning can be conducted after location planning, a joint design of location planning and frequency planning can achieve better performance and however is much more challenging. We can extend the current research of Chapter 5 by relaxing the assumption of frequency reuse factor of one. In other words, another design dimension can be considered together with location planning to further increase the system performance in terms of wireless bandwidth utilization.

- Placement in Wireless Sensor Networks (WSNs) under CT.

The wireless sensor networks are application-oriented, such as security surveillance, environment and weather observation, workplace and military applications. The network lifetime and/or connectivity are always the main concerns for WSNs. Besides, the robustness and resilience to point of failure are also very important especially in the security-aware applications. For instance, the WSN for security monitoring may require that each point is covered by at least two sensors such that once any sensor is compromised, another sensor still can operate. CT can be applied in wireless sensor networks which operate on batteries. By taking advantage of the CT technology, energy efficiency can be improved and thus the network lifetime can be increased. An optimal placement of a WSN under CT can not only improve the fault tolerant and robustness, but also can increase the network lifetime.

- Placement in Roadside-to-Vehicle Telematics Network.

In vehicular networks, some wireless devices such as routers or sensors are deployed on the roadside, acting as APs. The roles of APs are often two-fold: one is to provide connectivity to travelers that require Internet access services, the other is to monitor the traffic situation and react to accidents quickly by broadcasting alarm messages. The APs are connected to *gateways* which have direct links to the Internet. Thus, a

vehicular network constitute a three-tier architecture that consists of gateways, roadside APs and vehicular terminals. Then, two placement problems arise for infrastructure deployment, that is, AP placement and gateway deployment. The goal of AP placement is to cover the whole road so that every vehicle can be associated with at least an AP. And for gateway placement, since making every AP connected to a gateway is neither unnecessary nor practical most of the time. Instead, we may want to place just some gateways in the network. Thus, the problem is how to optimally place the gateways in terms of some specific requirements, such as minimal total power consumption, or minimal average hop numbers from APs to gateways. Due to the server fading effects caused by vehicle movements, similarly, CT can be applied to improve the transmission rate and reliability by virtue of spatial diversity gain. Therefore, we can also follow the methodology of this research and extend the optimization framework to the AP and gateway placement in vehicular networks under CT.

- Location Planning in Relay-based Integrated PON-WCN.

In relay-based integrated PON-WCN architecture, RSs can be introduced such that the network entities in the tree/PMP topology compromises an OLT, splitters, ONU-BSs, RSs and SSs. With RSs in the integrated architecture, the single-cell CT technology can also be applied together with the multi-CT technology (we have studied in Chapter 5). Even better performance gains in terms of capacity/throughput and reliability enhancement can be achieved by properly deploying all the network equipments (e.g., ONU-BSs and RSs). Thus, for a network provider, given the financial budget of building a new BAN, maximizing the network capacity should be considered as the main objective. The inputs includes the geographical distribution of traffic demands of VIPs, the mobile service requirement, the attenuation factor, the shadowing margins within AOI, maximal transmission power of each ONU-BS and RS, and wireless

bandwidth upper bound, the cost parameters of each network entity in the relay-based integrated PON-WCN, etc. The output of interests may includes the optimal locations of ONU-BSs, splitters, and RSs as well as fibre layout, the association between ONU-BSs and RSs, the association between the serving RSs to VIPs/TPs, the resource (transmission power and wireless bandwidth) breakdown assignment at each ONU-BS and RS. The problem is much more challenging since it will at least incorporate the studies in Chapters 3-5 of this dissertation. Also, it is more challenging to develop a solution to deal with the huge search space due to the large number of decision variables. To solve the location planning in such a relay-based integrated PON-WCN, we may also follow the methodology by proposing a problem-specific mathematical formulation and a new solution approach.

6.3 Final Remarks

In this dissertation, we have shown that, given a set of network parameters, the proposed optimization framework can provide design guidelines for practical network deployment and cost estimations. The constructed broadband access networks are cost-effective by taking advantage of CT technology. In most situations, the location planning is conducted offline, and the calculations within several hours/days are applicable in reality. For emergent deployment scenarios, an efficient solution (e.g., the proposed fast heuristic algorithm in *Chapter 3*) is suited for timely adaption to traffic boost.

Appendix A

Appendix

A.1 Calculation of the Maximal Chord-Curve Distance D_{UB}.

The maximal chord-curve distance D_{UB}, given the values of ξ^k_{min} and ξ^k_{max}, is determined by the linear functions $H^k(\xi_k) = a\xi_k + b$ and $G^k(\xi_k) = a\xi_k + b_1$ (see Figure A.1), where the slope a, intercept b and b_1 are calculated as follows. For the line $H^k(\xi_k)$, a and b can be obtained by the coordinates of point $A(\xi^k_{min}, \frac{\rho_k}{\log(1+\xi^k_{min})})$ and point $B(\xi^k_{max}, \frac{\rho_k}{\log(1+\xi^k_{max})})$, both of which are on the line $H^k(\xi_k)$. Then,

$$a = \frac{f_B(\xi^k_{max}) - f_A(\xi^k_{min})}{\xi^k_{max} - \xi^k_{min}} = \frac{\rho_k}{\xi^k_{max} - \xi^k_{min}}[\frac{1}{\log(1+\xi^k_{min})} - \frac{1}{\log(1+\xi^k_{max})}] \tag{A.1}$$

$$b = \frac{\rho_k}{\log(1+\xi^k_{min})} - a\xi^k_{min} \tag{A.2}$$

The intercept of the line $G^k(\xi_k)$, i.e., b_1, is calculated based on point $C(x_c, \frac{\rho_k}{log(1+x_c)})$, where $C \in G^k(\xi_k) = a\xi_k + b_1$, therefore:

$$b_1 = \frac{\rho_k}{\log(1+x_c)} - ax_c \tag{A.3}$$

By observing $\triangle CEF$ and $\triangle CDE$, we yield $\sin\theta = \frac{|CF|}{|EF|} = \frac{|CD|}{|CE|}$. Then, $D_{UB} = |CD|$ is determined by

$$\frac{b - b_1}{c} = \frac{m}{\frac{b}{a} - \frac{b_1}{a}}.$$

Pythagoras's theorem implies for $c = |EF|$:

$$c = \sqrt{(b - b_1)^2 + (\frac{b}{a} - \frac{b_1}{a})^2} = (b - b_1)\sqrt{1 + a^{-2}} \tag{A.4}$$

Therefore, given the values of ξ^k_{min} and ξ^k_{max}, based on (A.1), (A.2), (A.3), and (A.4), the maximal chord-curve distance D_{UB} can be obtained by

$$D_{UB} = \frac{b - b_1}{\sqrt{1 + a^2}} \tag{A.5}$$

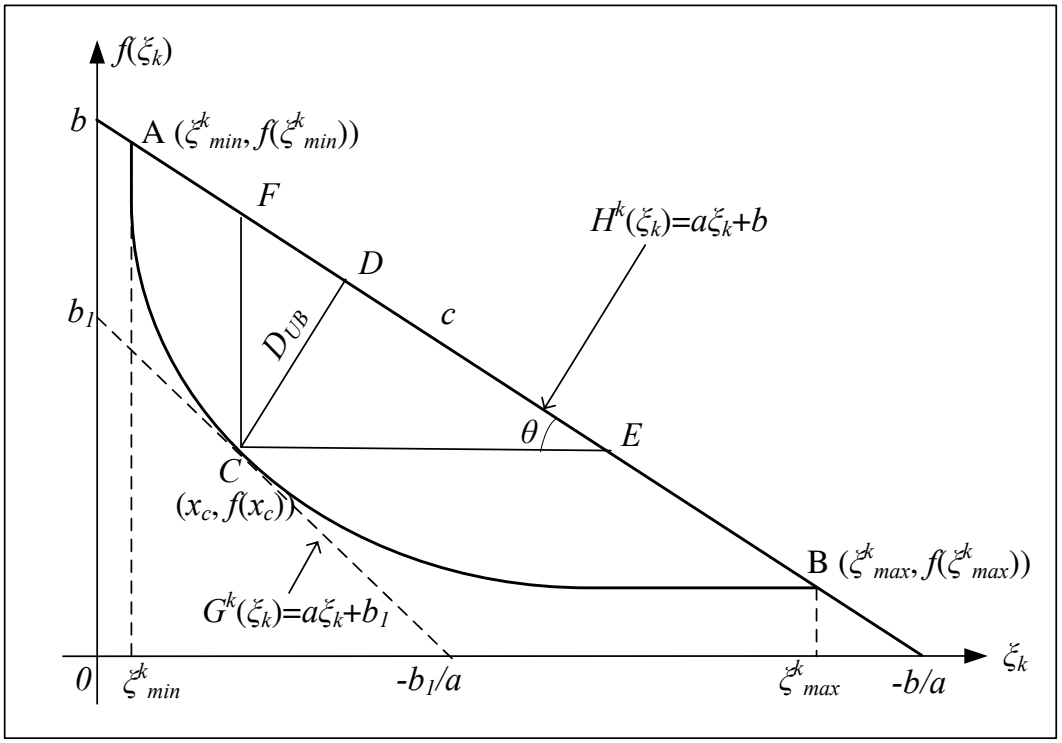


Figure A.1: Maximal chord-curve distance.

Bibliography

[1] T. Monath, N. K. Elnegaard, P. Cadro, D. Katsianis, and D. Varoutas, "Economics of fixed broadband access network strategies," *IEEE Commun. Mag.*, vol. 41, no. 9, pp. 132–139, Sep. 2003.

[2] M. Vrdoljak, S. I. Vrdojak, and G. Skugor, "Fixed-mobile convergence strategy: technologies and market opportunities," *IEEE Commun. Mag.*, vol. 38, no. 2, pp. 116–121, Feb. 2007.

[3] B. O. Obele and M. Kang, "Fixed mobile convergence: a self-aware QoS architecture for converging WiMAX and GEPON access networks," in *Proc. IEEE NGMAST*, Cardiff, Wales, UK, Sept. 2008, pp. 411 – 418.

[4] P. Beastall, "Bringing broadband to fixed mobile convergence," in *Proc. IET Wireless Broadband Conf.*, London, UK, Apr. 2006, pp. 1 – 16.

[5] T. Rokkas, D. Katsianis, D. Varoutas and T. Sphicopoulos, "Fixed mobile convergence for an integrated operator: a techno-economic study," in *Proc. IEEE PIMRC*, Athens, Greece, Sept. 2007, pp. 1 – 5.

[6] A. Sendonaris, E. Erkip and B. Aazhang, "User cooperation diversity – part I: system description," *IEEE Trans. Computers*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[7] L.-L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Trans. Information Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[8] H. Zhang and H. Dai, "Co-channel interference mitigation and cooperative processing in downlink multi-cell multiuser MIMO networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2004, no. 2, pp. 222–235, . 2004.

[9] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Information Theory*, vol. 50, pp. 748–767, May 2004.

[10] J. Yang and H. Lee, "An AHP decision model for facility location selection," *Facilities*, vol. 15, no. 9/10, pp. 241–254, Sep. 1997.

[11] M. R. Garey and D. S. Johnson, *Computers and intractability: a guide to the theory of NP-completeness.* New York: Freeman, 1979.

[12] J. Sydir, "Harmonized Contribution 802.16j (Mobile Multihop Relay) Usage Models," online, http://grouper.ieee.org/groups/802/16/relay/docs/80216j-06_015.pdf, 2006.

[13] B. Lin, P. H. Ho, L. L. Xie and X. Shen, "Optimal Relay Station Placement in IEEE 802.16j Networks," in *Proc. ACM IWCMC*, Honolulu, Hawaii, Aug. 12-16 2007.

[14] B. Lin, P. H. Ho, L. L. Xie and X. Shen, "Relay Station Placement in IEEE 802.16j Dual-Relay MMR Networks," in *Proc. IEEE ICC*, Beijing, China, May 19-23 2008.

[15] B. Lin, M. Mehrjoo, P. H. Ho, L. L. Xie and X. Shen , "Capacity enhancement with relay station placement in wireless cooperative networks ," in *Proc. IEEE WCNC*, Budapest Hungary, Apr. 2009.

[16] X. Qiu, P. Ossieur, J. Bauwelinck, Y. Yi, D. Verhulst, J. Vandewege, B. De Vos and P. Solina, "Development of GPON upstream physical-media-dependent prototypes," *Journal of Lightwave Tech.*, vol. 22, no. 11, pp. 2498 – 2508, Nov. 2004.

[17] M. Hajduczenia, H. J. A. Da Silva and P. Monteiro, "Development of 10 Gb/s EPON in IEEE 802.3av," *IEEE Commun. Mag.*, vol. 46, no. 7, pp. 40 – 47, Jul. 2008.

[18] G. Kramer and G. Pesavento, "Ethernet passive optical network (EPON): building a next-generation optical access network," *IEEE Commun. Mag.*, vol. 6, no. 5, pp. 66–73, Feb. 2000.

[19] G. Shen, R. S. Tucker and C. Chae, "Fixed mobile convergence architectures for broadband access: integration of EPON and WiMAX," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 44–50, Aug. 2007.

[20] B. Lin, P. H. Ho and X. Shen , "Network Planning for Next-Generation Metropolitan-Area Broadband Access under EPON-WiMAX Integration ," in *Proc. IEEE Globecom*, New Orleans, LA, USA, Nov. 30 - Dec. 4 2009.

[21] S. Sarkar, S. Dixit, and B. Mukherjee, "Hybrid wireless-optical broadband-access network (WOBAN): a review of relevant challenges," *IEEE Journal of Lightwave Tech.*, vol. 25, no. 11, pp. 3329–3340, Nov. 2007.

[22] Wong, N. Cheng, K. Balasubramanian, X. Zhu, M. Maier, and L. G. Kazovsky, "Hybrid architecture and integrated routing in a scalable optical-wireless access network," *IEEE Journal of Lightwave Technology*, vol. 25, no. 11, pp. 3443–3451, Nov. 2007.

[23] T. M. Cover and A. A. El. Gamal, "Capacity theorems for the Relay Channel."

[24] L.-L. Xie and P. R. Kumar, "Multi-source, multi-destination, multi-relay wireless networks," *IEEE Trans. Information Theory*, vol. 53, no. 10, pp. 3586–3595, Oct. 2007.

[25] M. Gastpar, G. Kramer and P. Gupta, " The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. IEEE ISIT*, Lausanne, Switzerland, Jul. 2002, p. 136.

[26] T. Himsoon, W. P. Siriwongpairat, Z. Han and K. J. R. Liu, "Lifetime maximization via cooperative nodes and relay deployment in wireless networks," *IEEE JASC*, vol. 25, no. 2, pp. 306 – 317, Feb. 2007.

[27] S. Kim, X. Wang and M. Madihian, "Optimal resource allocation in multi-hop OFDMA wireless networks with cooperative relay," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 1833 – 1838, May 2008.

[28] Y. Shi, S. Sharma, Y. T. Hou and S. Kompella , "Optimal relay assignment for cooperative communications ," in *ACM Mobihoc*, Hong Kong, China, May 2008, pp. 3–12.

[29] A. Bletsas, A. Khisti, D. Reed and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE JSAC*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[30] J. Cai, X. Shen, J. W. Mark and A. S. Alfa, "Semi-distributed user relaying algorithm for amplify-and-forward wireless relay networks," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 4, pp. 1348–1457, Apr. 2008.

[31] A. S. Ibrahim, A. K. Sadek, W. Su and K. J. R. Liu, "Cooperative communications with relay-selection: when to cooperate and whom to cooperate with?" *IEEE Trans. on Wireless Commun.*, vol. 7, no. 7, pp. 2814 – 2827, Jul. 2008.

[32] I. Krikidis, J. Thompson and N. Goertz, "A cross-layer approach for cooperative networks," *IEEE Trans. on Vehi. Tech.*, vol. 57, no. 5, pp. 3257 – 3263, Sept. 2008.

[33] J. N. Laneman, *Cooperative diversity in wireless networks: algorithms and architectures.* Ph.D. dissertation, MIT, Cambridge, MA , Aug. 2002.

[34] O. Somekh, O. Simeone, Y. B. Ness, and A. M. Haimovich, "Distributed multi-cell zero-forcing beamforming in cellular downlink channels," in *Proc. IEEE GLOBECOM*, San Francisco, California, USA, Nov. 2006, pp. 1–6.

[35] L. Shao and S. Roy, "Downlink multicell MIMO-OFDM: an architecture for next generation wireless networks," in *Proc. IEEE WCNC*, New Orleans, LA, USA, Mar. 2005, pp. 1120–1125.

[36] I. D. Garcia, K. Sakaguchi, and K. Araki, "Cell planning for cooperative transmission," in *Proc. IEEE WCNC*, Las Vegas, Nevada, USA, Mar. 2008, pp. 1769–1774.

[37] A. Tolli, M. Codreanu and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1428 – 1440, Apr. 2008.

[38] Y. Kim, H. Li and H. Liu , "Capacity of a multi-cell cooperative system," in *Proc. IEEE ISWCS*, Reykjavik, Iceland, Oct. 2008, pp. 314 – 318.

[39] M. Daskin, *Network and discrete location: models, algorithms and applications.* John Wiley and Sons Canada, Limited, 1 Edition, 1995.

[40] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 7–12, Mar. 1960.

[41] N. Mladenovic, J. Brimberg, P. Hansen, M. Perez and A. Jose, "The p-median problem: A survey of metaheuristic approaches," *European J. of Operational Research*, vol. 179, no. 3, pp. 927–939, Jun. 2007.

[42] B. C. Tansel, R. L. Francis and T. J. Lowe, "State of the artlocation on networks: a survey. part i: the p-center and p-median problems," *Management Science*, vol. 29, no. 4, pp. 482–497, Apr. 1983.

[43] M. S. Daskin, "A maximum expected covering location model: formulation, properties and heuristic solution," *Transportation Science*, vol. 17, no. 1, pp. 48–70, Feb. 1983.

[44] N. Christofides and S. Kormann, "A computational survey of methods for the set covering problem," *Management Science*, vol. 21, no. 5, pp. 591–599, Jan. 1975.

[45] Z. Drezner, *Facility location: a survey of applications and methods.* Springer Series in Operations Research and Financial Engineering, 1995.

[46] L. Ruan, D. Du, X. Hu, X. Jia, D. Li and Z. Sun, "Converter placement supporting broadcast in WDM optical networks," *IEEE Trans. Computers*, vol. 50, no. 7, pp. 750 – 758, Jul. 2001.

[47] X. H. Jia, D. Z. Du, X. D. Hu, H. J. Huang and D. Y. Li, "Placement of wavelength converters for minimal wavelength usage in WDM networks," in *Proc. IEEE INFOCOM*, vol. 3, Jun. 2002, pp. 1425–1431.

[48] X. Chu, B. Li and I. Chlamtac, "Wavelength converter placement under different RWA algorithms in wavelength-routed all-optical networks," *IEEE Trans. Computers*, vol. 51, no. 4, pp. 607 – 617, Apr. 2003.

[49] A. V. Tran, R. S. Tucker, N. L. Boland, "Amplifier placement methods for metropolitan WDM ring networks," *Journal of Lightwave Technology*, vol. 22, no. 11, pp. 2509 – 2522, Nov. 2004.

[50] M. L. Rocha,S. M. Rossi, M. R. X. Barros, L. Pezzolo, J. B. Rosolem, M. F. Oliveira, A. Paradisi, T. Kauppinen and A. Gavler, "Amplifier placement in metro-scaled wavelength-routed network," *Journal of Lightwave Technology*, vol. 39, no. 3, pp. 302 – 304, Feb. 2003.

[51] H. Lin and S. Wang, "Splitter placement in all-optical WDM networks," in *Proc. IEEE GLOBECOM*, vol. 1, St. Louis, USA, Nov 2005, p. 5.

[52] T. Hu and B. Zhao, "Splitter Placement problem on directed fiber trees," in *Proc. IEEE PDCAT*, Dunedin, New Zealand, Dec. 2005, pp. 225 – 228.

[53] E. Amaldi, A. Capone and F. Malucelli, "Planning UMTS base station location: Optimization models with power control and algorithm," *IEEE Wireless Commun.*, vol. 2, no. 5, pp. 939–952, Sep. 2003a.

[54] N. Weicker, G. Szabo, K. Weicker, and P. Widmayer, "Evolutionary multiobjective optimization for base station transmitter placement with frequency assignment," *IEEE Trans. on Evolutionary Computation*, vol. 7, no. 2, pp. 189 – 203, Apr. 2003.

[55] S. Baeg and T. Cho, "Transmission relay method for balanced energy depletion in wireless sensor networks using fuzzy logic," *Fuzzy Systems and Knowledge Discovery*, vol. 3614, pp. 998–1007, 2005.

[56] S. Lee, J. Yoo and T. Chung, "Distance-based energy efficient clustering for wireless sensor networks," in *Proc. IEEE Int. Conf. on Local Computer Networks*, Tampa, Florida, Nov. 2004, pp. 567 – 568.

[57] J. R. Sack and J. Urrutia, *Art gallery and illumination problems.* Handbook of Computational Geometry. Elsevier Science B.V., Amsterdam, The Netherlands,, 2000.

[58] K. Xu, Q. Wang, H. Hassanein and G. Takahara, "Optimal design of wireless sensor networks: minimum cost with lifetime constraints," in *Proc. IEEE WiMob*, Montreal, Canada, Agu. 2005.

[59] Q. Wang, K. Xu, H. Hassanein, and G. Takahara, "Minimum cost guaranteed lifetime design for heterogeneous wireless sensor networks," in *Proc. IEEE IPCCC*, Phoenix, Arizona, Apr. 2005, pp. 599–604.

[60] P. Maulin, R. Chandrasekaran and S. Venkatesan, "Energy efficient sensor, relay and base station placements, for coverage, connectivity and routing," in *Proc. IEEE IPCCC*, Phoenix, Arizona, Apr. 2005, pp. 581 – 586.

[61] A. So and B. Liang, "An efficient algorithm for the optimal placement of wireless extension points in rectilineal wireless local area networks," in *Proc. IEEE QShine*, Orlando Florida USA, Aug. 2005, pp. 25–33.

[62] A. So and B. Liang, "A lagrangian approach for the optimal placement of wireless relay nodes in wireless local area networks," in *Proc. Int. IFIP-TC6 Networking Conf.*, Coimbra, Portugal, May 2006, pp. 160–172.

[63] A. So and B. Liang, "Enhancing WLAN capacity by strategic placement of tetherless relay points," *IEEE Trans. Mobile Comput.*, vol. 6, no. 5, pp. 522–535, May 2007.

[64] R. Chandra, L. L. Qiu, K. Jain, and M. Mahdian, "Optimizing the placement of Internet TAPs in wireless neighborhood networks," in *Proc. IEEE ICNP*, Berlin, Germany, Oct. 2004, pp. 271–282.

[65] A. So and B. Liang, "Optimal placement of relay infrastructure in heterogeneous wireless mesh networks by bender's decomposition," in *Proc. IEEE QShine*, Waterloo, Canada, Aug. 2006.

[66] E. Amaldi, A. Capone, M. Cesana, F. Malucelli , " Optimization models for the radio planning of wireless mesh networks," in *IFIP Networking*, Atlanta, Georgia, USA, May 2007.

[67] E. Amaldi, A. Capone, M. Cesana, I. Filippini and F. Malucelli, "Optimization models and methods for planning wireless mesh networks," *Computer Networks*, vol. 52, no. 11, pp. 2159–2171, Aug. 2008.

[68] H. Sherali, C. Pendyala, and T. Rappaport, "Optimal location of transmitters for micro-cellular radio communication system design," *IEEE JASC*, vol. 14, no. 4, pp. 662–673, May 1996.

[69] Y. Chen and H. Kobayashi , "Signal strength based indoor geolocation ," in *Proc. IEEE ICC*, New York, US, Apr. 2002.

[70] M. Wright, " Optimization methods for base station placement in wireless applications," in *Proc. IEEE VTC*, Ottawa, Canada, May 1998.

[71] L. Nagy and L. Farkas, "Indoor base station location optimization using genetic algorithms ," in *Proc. IEEE PIMRC*, London, UK, Sept. 2000.

[72] R. Battiti, M. Brunato and A. Delai , " Optimal wireless access point placement for location-dependent services," in *technical report*, University of Trento, Italy, Oct. 2003.

[73] Z. Jia, J. Yu, G. Ellinas and G.-K. Chang, "Key enabling technologies for opticalCwireless networks: optical millimeter-wave generation, wavelength reuse, and architecture," *Journal of Lightwave Tech.*, vol. 25, no. 11, pp. 3452 – 3471, Nov. 2007.

[74] C. Bock, M. P. Thakur, T. Quinlan, S. E. M. Dudley and S. D. Walker, "Integration of wireless and optics, future trends on access networks," in *Proc. ICTON*, vol. 4, Athens, Greece, Jun. 2008, pp. 73 – 77.

[75] S. R. Chaudhry and H. S. AL-Raweshidy, "Application-controlled handover for heterogeneous multiple radios over fibre networks," *IET Communications*, vol. 2, no. 10, pp. 1239 – 1250, Nov. 2008.

[76] H. Al-Raweshidy and S. Komaki , *Radio over fiber technologies for mobile communications networks.* Artech House , 2002.

[77] A. J. Cooper, "Fiber-radio for the provision of cordless/mobile telephony services in the access network," *Electron. Lett.*, vol. 26, pp. 2054–2056, . 1990.

[78] W.-P. Lin, "A robust fiber-radio architecture for wavelength-division multiplexing ring-access networks," *Journal of Lightwave Tech.*, vol. 23, no. 9, pp. 2610–2620, Sep. 2005.

[79] H. Pfrommer, M. A. Piqueras, J. Herrera, V. Polo, A. Martinez, S. Karlsson, O. Kjebon, R. Schatz, Y. Yu, T. Tsegaye, C. P. Liu, C. H. Chuang, A. Enard, F. VanDijk, A. J. Seeds, and J. Marti, "Full-duplex DOCSIS/WirelessDOCSIS fiber-radio network employing packaged AFPM-based base-stations," *IEEE Photon. Tech. Lett.*, vol. 18, no. 406-408, p. 2, Jan. 2006.

[80] S. Ray, M. Medard and L. Zheng , " A SIMO fiber aided wireless network architecture," in *Proc. ISIT*, Seattle, WA, Jul. 2006, pp. 2904–2908.

[81] C. P. Liu, T. Ismail and A. J. Seeds, "Broadband access using wireless-over-fibre technologies," *Springer BT Technol*, vol. 24, no. 3, pp. 130–143, Jul. 2006.

[82] B. Y. Choi, J. Park and Z. L. Zhang , "Adaptive random sampling for traffic load measurement ," in *IEEE ICC*, vol. 3, Anchorage, Alaska, USA, May 2003.

[83] J. W. Mark and W. Zhuang, *Wireless Communications and Networking.* Prentice Hall, 2003.

[84] ILOG CPLEX 10.0/11.0, *CPLEX Optimization Inc.* , 2006.

[85] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning, reading.* MA: Addison-Wesley, 1989.

[86] M. Kojima, N. Megiddo and S. Mizuno, "Theoretical convergence of large-step-primal-dual interior point algorithms for linear programming," *Mathematical Programming*, pp. 1–21, 1993.

[87] B. Lin and P. H. Ho, "Dimensioning and Location Planning for Wireless Networks under Multi-level Cooperative Relaying," in *IFIP Networking*, Aachen, Germany, May 11-15 2009, pp. 207–219.

[88] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE JSAC*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[89] K. Ohara, V. J. Hernandez, Y. Du, Z. Ding, S. J. B. Yoo and Y. Horiuchi, "Resiliency of OCDM-PON against near-far problem," in *Technical Digest of IEEE/OSA Optical Fiber Communication Conference*, no. OMO3, 2007.

[90] IEEE 802.16e/D12, *Air interface for fixed and mobile broadband wireless access systems.* IEEE Standard for local and metropolitan area networks, 2005.