DEMAND-BASED NETWORK PLANNING FOR WLANS

by

Chutima Prommak

B.S. in Electrical Engineering, Khon Kaen University, 1992

M.S. in Interdisciplinary Telecommunications, University of Colorado at Boulder, 1998

Submitted to the Graduate Faculty of

School of Information Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2004

UNIVERSITY OF PITTSBURGH

SCHOOL OF INFORMATION SCIENCES

This dissertation was presented

by

Chutima Prommak

It was defended on

June 9, 2004

and approved by

Dr. Prashant Krishnamurthy, Assistant Professor,
School of Information Sciences, University of Pittsburgh

Dr. Bryan Norman, Associate Professor,
Industrial Engineering, University of Pittsburgh

Dr. Sumit Roy, Professor
Electrical Engineering, University of Washington

Dissertation Co-Director: Dr. Joseph Kabara, Assistant Professor,
School of Information Sciences, University of Pittsburgh

Dissertation Co-Director: Dr. David Tipper, Associate Professor,
School of Information Sciences, University of Pittsburgh

To my father, Sungkom Prommak,

whose strength, perseverance and integrity have been an essential motivation to me

DEMAND-BASED NETWORK PLANNING FOR WLANS

Chutima Prommak, Ph.D.

University of Pittsburgh, 2004

The explosive recent growth in Wireless Local Area Network (WLAN) deployment has generated considerable interest among network designers. Previous design approaches have mostly focused on coverage based optimization or the application of trial and error strategies. These only ensure that adequate signal strength is maintained in the intended service area. WLAN service environments, however, require a network designed to provide not only radio coverage but also adequate capacity (data rate) across the service area so that it can carry traffic load from a large number of users with certain Quality of Service (QoS) requirements. Thus, current design techniques are insufficient to provide data communication services to WLAN users.

In this dissertation, a novel approach to the WLAN design problem is proposed that takes into account user population density in the service area, traffic demand characteristics and the structure of the service area. The resulting demand-based WLAN design results in a network that provides adequate radio signal coverage and the required data rate capacity to serve expected user traffic demand in the service region. The demand-based WLAN design model is formulated as a Constraint Satisfaction Problem (CSP). An efficient heuristic solution technique is developed to solve the CSP network design problem in reasonable computational time. The solution provides the number of access points required and the parameters of each access point, including location, frequency channel, and power level.

Extensive numerical studies have been reported for various service scenarios ranging from a single floor with small and large service areas to a multiple floor design to a design that includes outside areas. The results of these studies illustrate that the demand-based WLAN design approach is more appropriate for the design of the WLAN systems than are existing coverage based design approaches. Additionally, extensive sensitivity analysis was conducted to study the effects of user activity level (traffic load), shadow fading, and the use of different path loss models in network design.

ACKNOWLEDGMENTS

First, I would like to thank God for providing me with the opportunity to experience, learn, and grow in the United State. God has provided me with everything I needed and helped me finish this work. I thank God for all.

I would like to express my gratitude to my advisors, Dr. David Tipper and Dr. Joseph Kabara. I appreciate their patience in working with me and waiting for me to learn and grow professionally. I highly appreciate their guidance and encouragement throughout the course of my research work. Additionally, I would like to thank my dissertation committee members. I would like to thank Dr. Prashant Krishnamurthy for his constructive comments and encouragement. My deep appreciation and gratitude go to Dr. Bryan Norman. His excellent lectures on heuristic algorithms and network optimizations have inspired and provided me with guidelines to develop the heuristic solution technique presented in this dissertation. I am highly indebted for his attention during long discussions over research problems and his invaluable comments and suggestions. I also would like to thank Dr. Sumit Roy for his constructive comments and suggestions that helped improve my research work.

Most certainly, I would like to acknowledge the following people for their profound insight and invaluable advice that they have provided during the course of my graduate study and dissertation research work. I would like to thank Dr. Rangsan Tongta, Dr. Hukju Kim, Dr. Jinseok Song, Dr. Joellen Popma, Dr. Wasu Chaopanon, Dr. Pramin

Artrit, Dr. Patchrawat Uthaisombut, Dr. Chalermpol Charnsripinyo, Dr. Peerapon Siripongwutikorn, Dr. Wipawee Thammapornphilas, Dr. Cholyeun Hongpaisanwiwat, Dr. Yu Liu, Dr. Siripun Sanguansintukual, Dr. Siripen Larpkiattaworn, and Dr. Kesorn Pechrach.

I would also like to thank my fellow Ph.D. students for generating an active research and study environment, and arranging recreation activities allowing me to recharge my energy and remain balanced. I thank Wiklom Teerapabkajorndet, Tanapat Anusas-Amornkul, Wasan Pattara-Atikom, Kamol Kaemarungsi, Pongsak Prasithsangaree, Gywn Chatranon, Chatree Sangpachatanaruk, Koonlachat Meesublak, Anotai Srikitja, Pantipa Phayak, Jumpol Polvichai, Rojana Leecharoen, Poonsiri Ariyawongrat, Marut Buranarach, and Ratchata Peachavanish. (At the time of this writing, they are finishing up their dissertation work and are Dr. soon to be.)

I also wish to acknowledge the wonderful staff at the School of Information Science, University of Pittsburgh. In particular, I would like to thank Arleen Girty, Regina Premozic, Theresa Benedek, Tricia McFadden, and Mark Steggert for helping me in many ways during the time of my study and research here.

Especially, I would like to thank the Royal Thai government for their financial support of my graduate study. I also would like to acknowledge the Telecommunications Program at the School of Information Science, University of Pittsburgh for providing to me financial support during the last two years of dissertation work. I would like to thank Suranaree University of Technology, Thailand for allowing me time to finish my graduate study and dissertation work.

My special thanks go to Matthew Shtrahman for his support, understanding, and patient listening. His passion towards his research work has been a great source of inspiration to me to finish up my dissertation.

Last, I would like to thank my dear brothers, Natthapol, Vechapan, and Kamon, for their support and love. Most importantly, I would like to express my deepest gratitude to my father, Sungkom Prommak, and mother, Yeesoon Prommak. I am very grateful for their guidance, advice, and constant love throughout my life. I thank them for providing me with everything they have done so that I was spared from many hardships and difficulties that they used to experience in their lives.

TABLE OF CONTENTS

ACKNOWLEDGMENTS .... **VII**

TABLE OF CONTENTS .... **X**

LIST OF TABLES .... **XVI**

LIST OF FIGURES .... **XVII**

1.0 INTRODUCTION **1**

1.1 WLAN OVERVIEW 2

1.2 ISSUES IN WLAN DESIGN 4

1.2.1 Network Capacity and User Density *4*

1.2.2 Coverage *5*

1.2.3 Frequency Channel Assignment *6*

1.2.4 Structure of Service Areas *7*

1.3 MOTIVATION 8

1.4 RESEARCH STATEMENT 9

1.5 RESEARCH CONTRIBUTIONS 11

1.6 OUTLINE OF DISSERTATION 11

2.0 LITERATURE REVIEW **12**

2.1 INDOOR WIRELESS NETWORK DESIGN 12

2.1.1 General Approaches to WLAN System Design *12*

2.1.2 Optimal Indoor Wireless Network Design *17*

2.2 CELLULAR NETWORK DESIGN 22

2.3 DEMAND NODE REPRESENTATION FOR TRAFFIC DEMAND IN CELLULAR NETWORK DESIGNS 24

2.3.1 Uniform-Size Demand Node 25

2.3.2 Variable-Size Demand Node 25

2.4 NETWORK USAGE CHARACTERISTIC IN WLANS 26

2.5 QUALITY OF SERVICE IN WLANS 28

2.6 CONSTRAINT SATISFACTION PROBLEM 29

2.7 PATH LOSS MODELS FOR INDOOR RADIO PROPAGATION 31

2.7.1 Single Floor Models 31

2.7.2 Multi-Floor Models 32

2.8 ANTENNA PATTERNS 33

2.8.1 Antenna Gain Equations 34

2.8.2 Antenna Patterns 35

2.9 SUMMARY 40

3.0 DEMAND-BASED WLAN DESIGN: MODEL AND PROBLEM FORMULATION 42

3.1 PROBLEM DEFINITION 43

3.2 NETWORK DESIGN CRITERIA 44

3.2.1 Radio Signal Coverage Requirement 44

3.2.2 User Data Rate Capacity Requirement 44

3.3 DEMAND NODE REPRESENTATION FOR WLAN DESIGN 45

3.3.1 Definition of a Demand Node for WLAN Design 46

3.3.2 WLAN Usage Characteristics and User Activity Level 48

3.4 DEMAND-BASED WLAN DESIGN MODEL 51

3.4.1 Definition 51

3.4.1.1 Signal Test Points 51

3.4.1.2 Demand Nodes 52

3.4.1.3 Network Configuration 52

3.4.2 Notation *53*

3.4.3 Static Parameters *55*

3.4.4 Dynamic Parameters *56*

3.4.5 Constraint Satisfaction Problem Formulation *62*

3.5 SOLUTION TECHNIQUE FOR THE DEMAND-BASED WLAN DESIGN 65

3.6 SUMMARY 66

4.0 A HEURISTIC SOLUTION TECHNIQUE FOR DEMAND-BASED WLAN DESIGN 67

4.1 A FRAMEWORK FOR SOLVING THE WLAN CSP 67

4.2 CONSTRUCTION PHASE 71

4.2.1 Area Covering Heuristic (ACH) *71*

4.2.2 Demand Clustering Heuristic (DCH) *75*

4.2.3 Example of the Construction Phase *76*

4.3 FREQUENCY CHANNEL ASSIGNMENT PHASE 79

4.3.1 Generating Channel Separation Constraints *80*

4.3.2 Assigning Frequency Channels to APs *83*

4.3.2.1 Overview of Simulated Annealing 84

4.3.2.2 Simulated Annealing Implementation for WLAN Frequency Channel Assignment 85

4.3.2.3 Procedures of the FCA SA Algorithm 88

4.3.2.4 Static Parameters in FCA Phase 90

4.4 CONSTRAINT VIOLATION REDUCTION PHASE 90

4.4.1 Overview of Tabu Search.. 91

4.4.2 Overall Structure of Tabu Search Procedures Implemented for the CVR Phase 93

4.4.3 Evaluation of Network Configuration.. 96

4.4.4 Constraint Violation Information 98

4.4.5 Neighborhood Structure and Move Operators 100

4.4.6 Short Term Memory 114

4.4.7 Long Term Memory.. 116

4.4.8 Aspiration criterion... 116

4.4.9 Stopping Criteria for CVR Phase 116

4.4.10 Static Parameters in the CVR Phase 117

4.5 INTENSIFICATION PHASE 118

4.6 ADD-AP PHASE 118

4.7 SUMMARY 119

5.0 SENSITIVITY ANALYSIS AND EFFECTIVENESS OF THE HEURISTIC SOLUTION TECHNIQUE 121

5.1 PARAMETERS IN THE DEMAND-BASED WLAN DESIGN METHODOLOGY 121

5.2 SENSITIVITY ANALYSIS OF STATIC PARAMETERS IN THE FCA PHASE 122

5.2.1 Parameter Settings for the Sensitivity Analysis 122

5.2.2 Test Networks.. 124

5.2.3 Numerical Results and Discussion... 128

5.2.3.1 Efficiency of the Simulated Annealing Method in the FCA Phase 129

5.2.3.2 Effects of P_{a_init} on the FCA Phase 132

5.2.3.3 Effects of φ on the FCA Phase 133

5.2.3.4 Effects of stopping criteria on the FCA phase 134

5.2.4 Parameter Setting of the Simulated Annealing Method in the FCA Phase ... 135

5.3 SENSITIVITY ANALYSIS ON STATIC PARAMETERS IN THE CVR PHASE 136

5.3.1 Parameter Settings for the Sensitivity Analysis 137

5.3.2 Test Scenarios 138

5.3.3 Numerical Results and Discussion.. 140

5.3.3.1 Effects of Tabu Tenure Length on the CVR Phase................................ 140

5.3.3.2 Effect of Max_{iter_CVR} on the CVR Phase... 142

5.3.3.3 Effect of the UA Size on the CVR Phase 143

5.3.3.4 Effect of Step size (λ) on the CVR Phase... 146

5.3.4 Parameter Setting for the CVR Phase 147

5.4 EFFECTIVENESS OF THE HEURISTIC SOLUTION TECHNIQUE 147

5.4.1 Exhaustive Enumeration Test 147

5.4.2 Comparison with other Solution Techniques... 149

5.5 SUMMARY... 151

6.0 NUMERICAL EXPERIMENTS 153

6.1 SINGLE FLOOR WLAN DESIGNS... 153

6.1.1 Experiment Settings 154

6.1.2 Numerical Results 158

6.2 MULTI-FLOOR WLAN DESIGN.. 162

6.2.1 Experimental Settings 162

6.2.2 Numerical Results 164

6.3 EXTEND INDOOR WLAN DESIGN TO A COURTYARD AREA 167

6.3.1 Experiment Settings 167

6.3.2 Numerical Results 169

6.4 EFFECTS OF USER ACTIVITY LEVEL 171
6.4.1 Experiment Settings 171
6.4.2 Numerical Results and Discussion........ 173
6.4.2.1 Effects of User Activity Level on Network Configurations 173
6.4.2.2 Comparative Performance Evaluation: Demand-based WLAN Design vs. Coverage-based WLAN Design 181
6.5 EFFECTS OF SHADOW FADING........ 185
6.5.1 Experimental Settings 186
6.5.2 Numerical Results and Discussions........ 186
6.6 EFFECTS OF PATH LOSS MODELS USED IN THE WLAN DESIGN 189
6.6.1 Experiment Settings 189
6.6.2 Numerical Results and Discussions........ 191
6.6.2.1 Is the design sensitive to path loss models?........ 191
6.6.2.2 Effect of Using Different Path Loss Models on the Resulting Network Configurations........ 192
6.7 SUMMARY........ 201
7.0 CONCLUSIONS, CONTRIBUTIONS, LIMITATIONS, AND FUTURE WORK 203
7.1 CONCLUSIONS AND CONTRIBUTIONS 203
7.2 LIMITATIONS AND FUTURE WORK 205
APPENDIX A 207
APPENDIX B 208
BIBLIOGRAPHY 217

LIST OF TABLES

Table 3.1 Static parameters 55

Table 3.2 Dynamic parameters 56

Table 5.1 Network usage characteristics 128

Table 5.2 Sensitivity analysis of control parameters in FCA phase: Network 1 130

Table 5.3 Sensitivity analysis of control parameters in FCA phase: Network 2 131

Table 5.4 Network usage characteristics 139

Table 5.5 Effects of tabu tenure on the CVR process 142

Table 5.6 Effects of Max_{iter_CVR} on the CVR process 143

Table 5.7 Effects of UA size on the efficiency of the CVR process 144

Table 5.8 Effects of step size (λ) on the CVR process 146

Table 5.9 Summary of the exhaustive enumeration test 149

Table 5.10 Comparison of solution techniques for the demand-based WLAN design problem 152

Table 6.1 Network usage characteristics 156

Table 6.2 Summary of parameters used in the WLAN design 157

Table 6.3 Parameters used in sensitivity analysis to study the effects of user activity level 173

Table 6.4 Comparative performance evaluation of WLAN configurations for SIS4 184

Table 6.5 Comparative performance evaluation of WLAN configurations for HL1 184

Table 6.6 Shadow fading and corresponding fading margins 186

Table 6.7 Initial network configuration 187

Table 6.8 Resulting WLAN configurations 188

LIST OF FIGURES

Figure 1.1 Infrastructure network 3

Figure 1.2 Frequency spectrum allocation for IEEE 802.11b and 802.11g........ 6

Figure 2.1 Idealized AP coverage [14] 15

Figure 2.2 APs' initial locations [14]........ 15

Figure 2.3 Uniform size demand node [36] 25

Figure 2.4 Variable-size demand node [39]........ 26

Figure 2.5 Spherical coordinate system [57] 34

Figure 2.6 Illustration of directivity........ 35

Figure 2.7 Omnidirectioanal antenna pattern [59] 36

Figure 2.8 Antenna pattern of the elevation plane of the dipole omnidirectional antenna [57] 37

Figure 2.9 Monopole antennas [57] 38

Figure 2.10 Antenna pattern of the elevation plane of the monopole omnidirectional antenna [57] 39

Figure 2.11 Antenna pattern of the patch directional antenna [57] 40

Figure 3.1 User distribution during different times of day 47

Figure 3.2 User population representation for WLAN design........ 48

Figure 3.3 Structure of a successful transmission time........ 62

Figure 4.1 Framework of the solution technique 70

Figure 4.2 Coverage Square (CS) estimating an AP's coverage area........ 72

Figure 4.3 Construction phase procedures........ 78

Figure 4.4 Pseudo code for generation of the channel separation matrix 81

Figure 4.6 The simulated annealing procedure for frequency channel assignments 89

Figure 4.7 Structure of Tabu search procedure implemented for the CVR phase........... 95

Figure 4.8 Examples of constraint violation information management......................... 100

Figure 4.9 Direction to move AP for conservative move operations 105

Figure 4.10 Example where STPs in a particular UA obtain their highest SIR levels from different APs.. 106

Figure 4.11 Short term memory: Direction attributes... 115

Figure 4.12 Outline of the intensification phase... 118

Figure 4.13 Add-AP procedure.. 119

Figure 5.1 Floor plans of the School of Information Science building 125

Figure 5.2 Effect of P_{a_init} on the computational time of the simulated annealing method 133

Figure 5.3 Floor-plan and demand node distribution for the first floor of Hillman Library (HL1) 139

Figure 5.4 Examples of problems in small or large UA size 145

Figure 6.1 Floor plans and demand node distribution of SIS4 and HL1 155

Figure 6.2 Grid point resolution of the small service area (SIS4) 158

Figure 6.3 WLAN network configuration for a small service area (SIS4).................... 160

Figure 6.4 WLAN network configuration for a large service area (HL1) 161

Figure 6.5 Floor plans of the fourth and fifth floors of the School of Information Science building.. 163

Figure 6.6 Signal coverage from APs located on different floors 165

Figure 6.7 WLAN configuration of the 4^{th} and 5^{th} floors of the SIS building............... 166

Figure 6.8 Layout of service area and demand node distribution of Hillman Library and the courtyard area next to the building 168

Figure 6.9 WLAN configuration for Hillman Library and the courtyard next to the building.. 170

Figure 6.10 Basic service areas and WLAN configuration of SIS4, low user activity level (scenario 1) 174

Figure 6.11 Basic service area and WLAN configuration of SIS4, medium user activity level (scenario 2) 175

Figure 6.12 Basic service areas and WLAN configuration of SIS4, high user activity level (scenario 3) 176

Figure 6.13 Basic service areas and WLAN configuration of HL1, low user activity level (scenario 4) 178

Figure 6.14 Basic service areas and WLAN configuration of HL1, medium user activity level (scenario 5) 179

Figure 6.15 Basic service areas and WLAN configuration of HL1, high user activity level (scenario 6) 180

Figure 6.16 Locations of APs in WLAN configurations obtained for different values of shadow fading.. 188

Figure 6.17 Comparison of WLAN configurations for SIS4 with low network traffic (scenario 1) 195

Figure 6.18 Comparison of WLAN configurations for SIS4 with medium network traffic (scenario 2) 196

Figure 6.19 Comparison of WLAN configurations for SIS4 with high network traffic (scenario 3) 197

Figure 6.20 Comparison of WLAN configurations for HL1 with low network traffic (scenario 4) 198

Figure 6.21 Comparison of WLAN configurations for HL1 with medium network traffic (scenario5) 199

Figure 6.22 Comparison of WLAN configurations for HL1 with high network traffic (scenario 6) 200

1.0 INTRODUCTION

Wireless local area network (WLANs) are experiencing tremendous growth and becoming increasingly popular. The use of unlicensed frequency spectrum ranges and inexpensive network equipment has facilitated the deployment of WLANs [1]. They are deployed in many places, such as, university campuses, corporate offices, health institutes, and public places like airports, coffee shops, and shopping malls [2]. WLAN cards are integrated into many of today's laptops and handheld computers, and they are optionally available for almost all computers. As WLAN-access devices have become cheaper, smaller and more powerful, the demand for WLAN services has increased, resulting in the phenomenal growth in the number of WLAN users [1-4].

As ongoing research extends the capabilities of WLANs to provide multimedia services at higher data rates, there will be an increase in the public's demand for and dependence on mobile data services. Providing users with Internet access with good Quality of Service (QoS) requires a network designed to provide sufficient aggregate data rate (bandwidth) to an area with a large number of users demanding high data rates. However, current WLAN planning approaches do not design networks for specific aggregate data rates. As the network service environment changes, there is a need to change the design strategy of large WLANs as well.

In this dissertation, we present a new approach to the network planning of large-scale WLANs. This chapter provides an introduction to the dissertation. Section 1.1 is an overview of WLANs. Section 1.2 discusses issues in WLAN design. Section 1.3 addresses motivations for this research. Section 1.4 details the research goals and problems studied in this dissertation. Section 1.5 provides an outline of the remainder of the dissertation.

1.1 WLAN Overview

The wireless LAN industry in North America has received great attention since the Federal Communications Commission (FCC) authorized the public use of the Industrial, Scientific, and Medical (ISM) bands (ranging from 902 MHz – 5.85 GHz) in 1985 [5]. Following this, the Institute for Electrical and Electronic Engineers (IEEE) 802.11 working group began the development of 802.11 WLAN standards [6]. The 802.11 standard defines the Medium Access Control (MAC) and physical layer functionality for wireless connectivity in a local area network. The MAC protocol provides mechanisms for controlling access to the shared wireless medium. The 802.11 physical layer specifies radio transmission techniques. Recent standards approved by the IEEE include 802.11a [7], 802.11b [8], and 802.11g [9]. The network design model and solution technique presented in this dissertation are applicable to these WLAN standards.

In the 802.11 WLAN system architecture, a primary building block consists of a group of Basic Service Sets (BSSs), groups of wireless terminals within a contiguous radio coverage area. A coverage area within which wireless terminals are free to move around and yet still remain connected is called a Basic Service Area (BSA) [5]. The 802.11 architecture can be categorized into two network configurations: ad hoc networks and infrastructure networks. In an ad hoc network, two or more wireless terminals form an independent BSS in which they can communicate directly with any other terminals, but cannot connect to a wired infrastructure. Alternatively, an infrastructure network employs an access point (AP) for central control of the communication between wireless terminals participating in a BSS. The AP provides a connection point to a wired network infrastructure such as an Ethernet LAN (802.3), allowing wireless terminals to exchange data packets with outside networks such as the Internet. BSSs operating in infrastructure mode do not allow wireless terminals to communicate directly with other terminals in the BSS. All data packets must be relayed through an AP and each AP can cover a service area ranging from 20 to 300 meters in radius [8]. For large service regions, a cellular architecture with multiple BSSs can be used in which the APs are interconnected via a wired distribution infrastructure to form a single system called an Extended Service Set (ESS). Figure 1.1 illustrates an ESS where three BSSs exist. Note that some of BSSs in the ESS can overlap so that there are no holes in the service region, providing greater capacity in areas with high user density and providing fault tolerance in areas requiring high reliability. In this dissertation, we consider the design of infrastructure-based WLANs. We aim to layout BSSs to cover a target region and, if possible, carry all traffic demand from users in the service area.

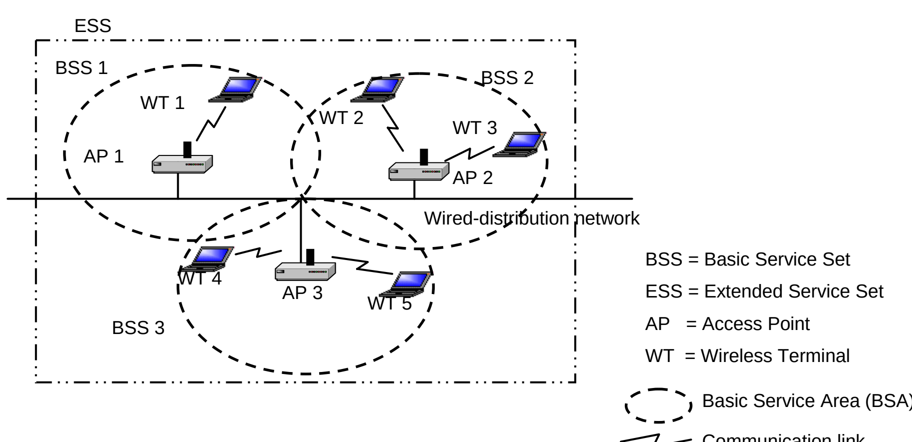


Figure 1.1 Infrastructure network

1.2 Issues in WLAN Design

Special concerns of WLAN design arise because of the unique characteristics of WLANs and service scenarios. In this section we discuss several issues that must be considered in the design of WLANs. They include network capacity and user density, coverage area, frequency assignment and structure of service areas.

1.2.1 Network Capacity and User Density

Wireless terminals in a BSS rely on a common (broadcast) transmission medium. Only one terminal can occupy the medium at a time. If multiple terminals simultaneously transmit, a collision may occur and the signal could be corrupted. The IEEE 802.11 standard specifies a Medium Access Control (MAC) protocol, called Carrier Send Multiple Access/ Collision Avoidance (CSMA/CA), to coordinate wireless terminal transmission in the BSS. This coordination is achieved by means of control information. This information is carried explicitly by control messages traveling in the medium (i.e. ACK messages) and can be provided implicitly by the medium itself through the use of a carrier-sensing mechanism before each transmission to check if the channel is either active or idle.

Control messages and message retransmission due to collisions consumes medium bandwidth. They are overhead required by a MAC protocol that coordinate wireless terminal transmissions. Although the 802.11b standard specifies that the AP can support channel data rates of 11 Mbps, BSS's actual capacity, defined as the fraction of channel bandwidth used by successfully transmitted messages, is less than 11 Mbps [10]. The practical throughput capacity decreases as the number of users (wireless terminals) associating with a particular AP increases [10, 11].

Providing sufficient data rate capacity is the first step towards any type of QoS support. Many data services (e.g. IP telephony, video conferencing, and multimedia applications) require that networks provide a specified average data rate. In addition to adequate signal strength, these applications require a guarantee of access channel capacity. Since each AP can provide only a limited data rate capacity and its throughput reduces as the number of users associated to it increases, a sufficient number of APs must be provided to support all traffic demand. However, one cannot over-deploy APs due to the limitation of frequency channels and interference problems among co-channel APs.

1.2.2 Coverage

A Basic Service Area (BSA) is the service coverage area of an AP. It is an area within which the received signal strength and the Signal to Interference Ratio (SIR) level are sufficient to allow data communication between an AP and wireless terminals to take place. The size of a BSA varies with the AP's power level and the radio propagation environment [12]. The AP's power level not only determines the signal strength received within the BSA but also affects interference and SIR level. The received signal strength at a particular location can be predicted using path loss models. In these models, the received signal level is estimated as a function of the distance and the radio propagation environment between an AP and a receiver [12]. Thus, network planning of WLANs must determine the appropriate power levels of APs in order to provide a specified signal strength while maintaining sufficiently low levels of interference in the service area.

1.2.3 Frequency Channel Assignment

The IEEE 802.11b and 802.11g operate at the 2.4 GHz ISM band while the IEEE 802.11a operates at the 5 GHz band. In North America, 83.5 MHz bandwidth is available in the 2.4 – 2.4835 GHz band while 300 MHz bandwidth is allocated in the 5.15 – 5.35 MHz (lower-band) and 5.725 – 5.825 MHz (upper-band). The 802.11 standard divides the 2.4 GHz band into eleven channels with center frequencies located 5 MHz apart as shown in Figure 1.2. Each channel has a frequency bandwidth of 22 MHz. Among these eleven channels are three channels whose bandwidths do not overlap each other. Those channels are 1, 6 and 11, as there is a frequency space of 3 MHz between channels 1 and 6 as well as between channels 6 and 11. These three channels are called the non-overlapping channels, and they can be assigned to adjacent APs without interfering with each other. The remaining channels overlap with one of the three non-overlapping channels and are called the overlapping channels.

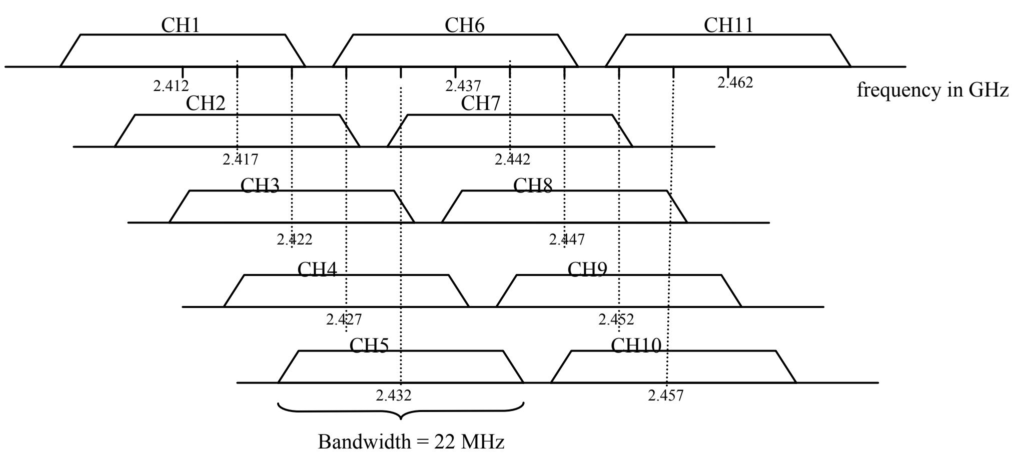


Figure 1.2 Frequency spectrum allocation for IEEE 802.11b and 802.11g

The 5 GHz lower-band of the 802.11a is divided into eight non-overlapping channels, while the 5 GHz upper-band is divided into four non-overlapping channels (each channel with 25 MHz bandwidth). There is no overlapping channel defined in the 5 GHz band of the 802.11 standard. The regulation specifies that the lower-band of the 5 GHz band be used for indoor WLANs and the upper-band for outdoor purposes only.

Since a limited number of channels exist in the available frequency spectrum for an 802.11 WLAN, a large network deployment requires that all channels are used and some channels are reused. Reuse of frequency channels in neighboring BSAs can cause interferences in the service area. Thus, frequency channels of APs must be carefully assigned in large WLANs.

1.2.4 Structure of Service Areas

Different service area environments pose particular problems for WLAN planning due to differences in building material and architecture [12]. In offices and classrooms, radio coverage can be difficult to achieve due to the high density of walls. In library study areas, large auditoriums and lecture halls, the service areas are rather open and there are fewer walls. In a multi-floor building, signals from adjacent floors complicate the WLAN design problem. When designing all types of networks, AP placement and frequency assignment must be properly designed considering differences in the physical structures of the service areas.

1.3 Motivation

Most existing indoor wireless network design approaches have focused on signal radio coverage objectives, including existing WLAN design approaches. Several general design approaches use software tools to facilitate the signal strength measurement process and simulate radio propagation in the service area, while existing optimization design approaches aim to ensure that an adequate signal strength is maintained in the intended service area and focus on the optimal access point/ base station placement problem with slightly different objective functions, generally varying only in their technique for solving the optimization problem.

The existing design approaches lack several key elements. First of all, traffic demand and user density are not considered. The coverage-based optimization approaches may appear sufficient for networks where user density is low and traffic load is light. However, we argue that they will be insufficient in future WLAN environments with higher user concentration and applications demanding increased data rates. In recent studies, Kabara [13] and Hills [14] separately discuss the need for consideration of capacity requirements in WLAN system design. However, the literature currently lacks a WLAN planning methodology that incorporates capacity requirements in the design process. Although traffic demand and user density issues have been studied in the context of cellular network design, those techniques cannot be directly applied to WLAN system design due to differences in the design objective, the nature of the traffic demand and medium access methods. Most cellular network designs aim to minimize the infrastructure cost (i.e., minimizing the number of base stations) while providing radio signal coverage. This is not the case for WLAN system design in which the infrastructure cost is insignificant compared to the system capacity due to the decreasing price of WLAN compliant devices. However, WLAN systems should be designed to provide an average required data rate to users in the service areas. A further different between cellular networks and WLANs is that the former carry voice traffic via Time Division Multiple Access (TDMA) or Code Division Multiple Access (CDMA) techniques while the latter WLANs typically carry Internet and multimedia applications via CSMA/CA techniques. The network design methodology needs to take into account these characteristics of WLANs.

Additionally, most of the existing WLAN design approaches are limited to small service coverage areas or to single floor service areas, two-dimensional (2-D) areas. WLANs have received a great deal of attention and been widely deployed in recent years. They are becoming practical in many settings, including university offices, cooperate offices, and campus buildings. A systematic approach to multi-floor WLAN service design is needed.

1.4 Research Statement

The goal of this dissertation is to develop a formal network design model and an efficient solution technique to the WLAN design problem. Unlike the coverage based design approach with objectives focusing on signal strength and interference level in the service area, we develop a demand-based WLAN design model which incorporates user density and expected traffic demand into the design process. Moreover, our network planning technique is capable of designing networks for three-dimensional (3-D, multi-floor) service regions, unlike current design approaches which are limited to small coverage areas and single floor service scenarios. The low cost of APs compared to the cost of the wireless devices with which they are communicating implies that it is unnecessary in the design of WLAN systems to minimize the number of APs, as is done in current optimization approaches. Thus, we formulate WLAN planning problems as constraint satisfaction problems (CSPs) with the objective of meeting the signal strength and interference level requirements as well as the expected traffic demand in the service region. Note that although the cost of APs are not a main concern here, providing too many APs in the service area leads to system performance degradation because of interference problems resulting from the limited available frequency spectrum. Therefore, only a sufficient number of APs should be placed in the network. Additionally, our CSP mathematical formulation incorporates an AP capacity analytical model in order to account for the effects of user density on AP capacity due to the nature of the medium access protocol CSMA/CA [10]. It also includes 3-D antenna pattern and multi-floor path loss models to capture propagation characteristics of the wireless signal in the multi-floor environment.

The goal of the presented network design for WLANs is to identify a sufficient number of APs and determine an efficient combination of the network parameters, including APs' locations, frequency channels, and power levels. Due to the complexity of the CSP network design models, a heuristic solution technique is developed to efficiently solve the demand-based WLAN design problem.

1.5 Research Contributions

In this dissertation, a novel demand-based WLAN design model has been developed and formulated as a CSP, taking into account both signal radio coverage requirements and average user data rate capacity requirements. Network usage characteristics of WLANs are accounted for by incorporating the correlation between users' network usage behavior and their locations into the CSP formulation. Furthermore, an efficient heuristic solution technique has been developed to solve the demand-based WLAN design problem. The framework of the developed demand-based WLAN design methodology is flexible and applicable to various network service environments, ranging from those with small, single floor service areas to those that are complex with multiple-floor service areas and those with a combination of indoor and outdoor service areas.

1.6 Outline of Dissertation

The remaining chapters of this dissertation are organized in the following manner:

Chapter 2 provides background information and a literature review of related research. Chapter 3 presents a new approach to WLAN design, describing the demand-based WLAN design model and presenting a network design problem formulation. Chapter 4 describes in detail the implementation of the heuristic solution technique for the demand-based WLAN CSP. Chapter 5 discusses a sensitivity analysis study on the key parameters of the heuristic solution technique. It also provides an effectiveness test of the heuristic solution technique. Chapter 6 details numerical studies on several aspects relevant to WLAN designs. Finally, chapter 7 concludes the dissertation and provides future research directions.

2.0 LITERATURE REVIEW

This chapter serves as a review of wireless network design approaches and other issues related to WLAN design. Section 2.1 reviews the existing indoor wireless network design approaches. Section 2.2 discusses cellular network design. Section 2.3 presents the demand node representation for traffic demand in cellular network design. Section 2.4 discusses network usage characteristics in WLANs. Section 2.5 concerns QoS issues in a WLAN. Section 2.6 provides an overview of the constraint satisfaction problem. Section 2.7 presents the path loss models for indoor radio propagation. Section 2.8 provides an overview of the antenna radiation patterns. And finally section 2.9 provides a summary of the chapter.

2.1 Indoor Wireless Network Design

Development of methodologies that help with WLAN system design has been limited. The related literature is classified into two categories: general design approaches and optimization design approaches.

2.1.1 General Approaches to WLAN System Design

WLAN systems can provide high data rates to mobile computers. In the past, utilization of WLAN systems was sparse with only a few number of users. The traditional methods of designing WLANs have been based on trial and error and involved placing access points (APs) in buildings at opportunistic locations [14-18]. Then, radio coverage is approximated from signal measurement and radio propagation prediction. Typically, network designers tweak the actual locations of APs based on coverage estimates. Site survey tools have been developed to facilitate the radio signal measurement process and create signal coverage maps based on the measured data. For example, both the Nokia Site survey tool [19] and InFielder [20] have features that draw a coverage map of an installed AP by moving a wireless terminal around within a target region to measure signal strength communicating between the wireless terminal and the AP. A number of radio propagation simulators have also been developed and are available commercially. For example, ProMan [21], Predictor [22] and CINDOOR [23] all have features to create a coverage map for an AP placed at a certain location by using indoor path loss models to predict radio propagation characteristics and coverage areas. Using these tools to create coverage maps, network designers can arbitrarily install APs to provide signal coverage in the target region. However, this does not account for user density and traffic demand characteristics. Thus, some areas in the service area may become crowded with a large number of wireless terminals competing for the same channel, resulting in lower APs' throughput and inadequate data rate capacity for the users' application requirements.

The general floor-plan method described in "Guidelines for IEEE 802.11 cell planning" [16] provides rules of thumb for installing a WLAN system. It provides a means of estimating the radio characteristics and the number of APs required by simply classifying structures and building materials into three types: open (no physical obstruction in the signal path), semi-open (shoulder-height partitions of wood or synthetic material), and closed (floor-to-ceiling walls of brick and plaster). For each type, the diameter of a circle representing an AP's coverage area is given. To determine the number of required APs and their locations, the circles are arbitrary drawn to cover the service area. This general floor-plan method does not guarantee coverage in a building structure because no path loss model is used to compute radio propagation in the service area. Again, user population and traffic demand are not considered.

Consider WLAN design approaches employed at Carnegie Mellon University (CMU). The goal of the first phase of the design was to provide WLAN service in some of the campus buildings. The basic strategy was to assume both that alternate floors would not need APs since they would be covered from above and below and that there would be constant coverage per AP [17]. The design process was to examine the drawings of the building, estimate coverage, place the APs, make spot measurements of actual signal to noise ratios, and fill coverage holes with additional APs [17]. This nonsystematic process could result in the overlap of APs' coverage areas. Later, the network was redesigned and reinstalled with an 802.11 WLAN operating at 2.4 GHz ISM band. CMU network designers then determined AP locations and assigned one of the three non-overlapping channels (1, 6, or 11) to each of the APs [14]. When placing the APs, the initial locations were selected by using the coverage-oriented approach aiming to avoid coverage gaps in the service areas and to minimize coverage overlaps among APs. The coverage radius is averaged from signal strength measurements. An idealized AP coverage is developed as three coaxial cylinders, as shown in Figure 2.1 [14]. The middle cylinder with radius R represents coverage on the floor on which the AP is located. The upper and lower cylinders with radius R' represent coverage on the floors above and below the one on which the AP is located. APs are initially placed within a service area as shown in Figure 2.2a for a single-floor building and in Figure 2.2b for a multi-floor building. The AP locations are adjusted and relocated based on signal strength measurements. Coverage maps are created to visualize the whole service area and to assign frequency channels to each AP. We can see that in the second phase of design, the main objective is again to provide signal strength coverage. Traffic profiles and user density are not taken into account. Frequency channel assignment is done manually. No mathematical formulation has been developed for the design.

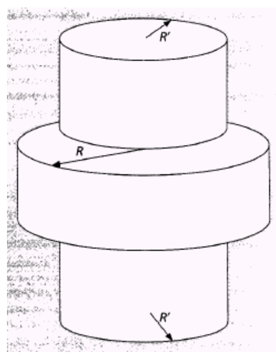
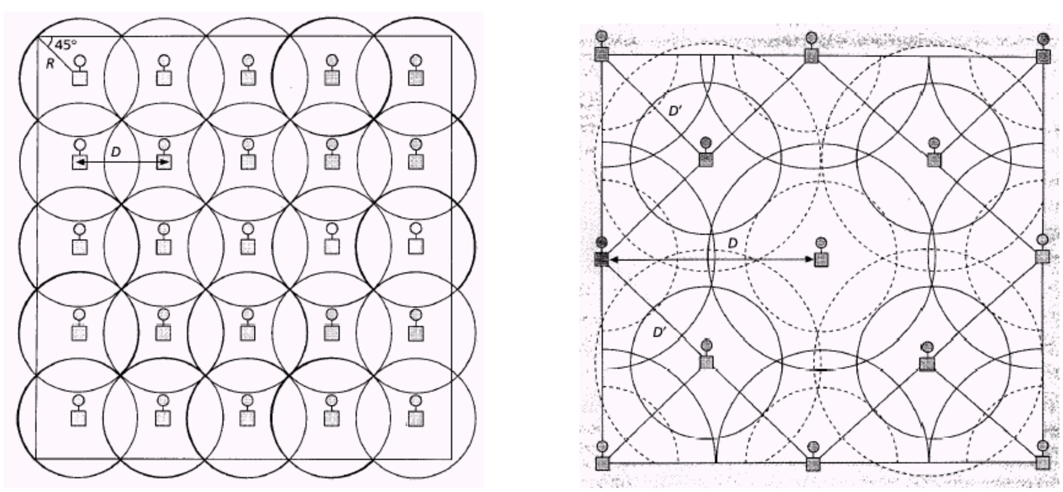


Figure 2.1 Idealized AP coverage [14]



a. single-floor area b. multi-floor building

Figure 2.2 APs' initial locations [14]

In a recent study [18], Unbehaun and Zander considered the WLAN system design for the 17 GHz band in an office environment. They focus on techniques to install APs to achieve signal coverage. They investigate the effects of different installation approaches, such as regular deployment and user deployment, on system performance. In regular deployment, APs are mounted on the ceiling in the center of a cell. This method requires additional wiring from the LAN outlets to the ceiling. Alternatively, in user deployment, APs are placed in opportunistic locations, such as sites close to the LAN outlets. This method aims to reduce installation costs by simplifying the wiring. The authors use a ray-tracing model to simulate indoor propagation of the office environment. They take into account shadowing effects caused by users' and others' bodies in the service areas, which will attenuate a number of rays arriving at the receiver. The quantitative results show that the authors' user-development approach yields acceptable signal coverage with the use of directional antennas. They suggest simple segmentation of the service area into equally sized cells and the placement of APs arbitrarily within the cells. However, they do not explain clearly how to segment the service area into cells, failing to address question about what the appropriate cell size is and how to compute the number of APs required. Further, their analysis is limited to the office environment. They do not consider traffic demand. They only aim to provide acceptable system coverage based on the path loss model computation.

Again in a later study, Unbehaun proposes a WLAN system operating at 17 GHz in an unlicensed band available in Europe [15]. The proposed system employs OFDM air interface. The author compares the frequency management based on two duplex schemes: Frequency Division Duplex (FDD) and Time Division Duplex (TDD). Three service scenarios with different building structures and user density are considered. Unbehaun aims to reduce installation cost by simplifying network planning and employing the user deployment approach or opportunistic AP placement [18]. Simulation results show that average data rate per user increases as additional APs are installed. However, although the user deployment approach might reduce the wiring cost, it can result in system performance degradation due to interference between neighboring APs.

2.1.2 Optimal Indoor Wireless Network Design

Some published works have investigated the design of indoor wireless networks, particularly WLANs, using optimization design approaches. These are reviewed as follows.

Adickes et al. [24] address the indoor wireless transceiver placement problem in a warehouse management system or a manufacturing execution system. To solve the design problem, they develop a heuristic optimization based upon a genetic algorithm. They aim to identify the optimal number and placement of transceivers in industrial facilities. Their design considers the radio signal coverage and average data rate capacity that a transceiver can achieve over the required coverage area. They employ the partitioning path loss model [12] to estimate a radio propagation environment. They use the Shannon limit for information capacity [25] to determine the data rate that can be transmitted across a medium. Coverage for the transceivers is initially modeled as a geometric problem, in which a number of circles are placed over an area in such a manner as to completely cover it with the minimum number of circles. The genetic algorithm then optimizes the transceivers' location according to radio signal and capacity requirements.

According to Shannon's formula, the data rate capacity is then computed based on the signal-to-noise ratio for a given medium bandwidth. However, the data rate capacity for WLANs cannot be simply computed from the signal strength. The MAC protocol and the users' traffic characteristics have an impact on the effective AP capacity. Thus, the transceiver placement approach of Adickes et al. [24] cannot be directly applied to the WLAN design.

Anderson and McGeehan [26] address the base station placement problem using the simulated annealing technique. They consider radio coverage and interference between multiple base stations, aiming to optimize signal levels in the service area by minimizing the mean square error of the difference between the receiver sensitivity and the power levels of each grid point in the service areas. They use a simplified empirical line-of-sight (LOS) path loss model [27] to approximate the radio environment. The empirical models are based on measurements and usually involve some kind of inverse function of distance, d, where the exponent of d is adjusted for the propagation conditions and whether the transmitter and receiver are in LOS or non-LOS conditions. Anderson and McGeehan's only objective in determining the location of the base stations is to improve the worst-case coverage situations. The overall performance (i.e., average received power for the entire service area) is not considered. In addition, the authors do not address the frequency channel nor the power level assignment for the base stations.

Sherali et al. [28] propose a mathematical design model to optimally place a single transmitter or a set of transmitters in a specified region and ensure that the signal strength at each point in the region is above a specified threshold. The problem is modeled by discretizing the radio coverage region into a grid of receiver locations and by specifying a function that estimates the path loss or signal attenuation for each receiver location, given a particular location for a transmitter that communicates with it. The authors use a site-specific propagation prediction model to evaluate the signal coverage [29], considering the forward link propagation case (i.e., the signal from the transmitter to the receiver). Sherali et al. formulate the design problem as a nonlinear programming problem with the objective of minimizing the measure of weighted path losses. The method employs a weighted convex combination of a minisum and a minimax objective. The minisum objective function aims to minimize the sum of path loss at each receiver to improve the average received power of the entire service area. The minimax objective function aims to minimize the highest path loss that violates the path loss threshold to ensure that even the worst receiver location is within the radio coverage. The convex combination of the two objective functions ensures good coverage over the entire service area, while maintaining an acceptable coverage at remotely located receivers. The authors developed nonlinear optimization solution techniques such as a conjugate gradient search approach to solve the problem. They state that "a region having a high density of users has a greater need for better coverage than a region where only a few users are likely to operate" [28]. While this is true in terms of providing the best signal quality for the highest number of users, signal quality alone is not sufficient without accompanying data rate capacity, especially in WLANs. Traffic load characteristics are not considered.

Stamatelos et al. [30] consider the transmitter placement problem for indoor wireless networks. Their goal is to determine locations and power levels of the transmitters. The design objective is to increase the number of users that can gain access to the network. To maximize user access capability, they try to provide sufficient signal strength across the service area with minimum interference in the areas where coverage of APs overlap. They formulate the optimization problem with an objective to minimize both a weighted sum of the area that is not covered and the sum of the areas that face excessive interference. The entire area to be covered is broken up into a rectangular grid and signal strength and interference are calculated within this grid. They assume the use of a single frequency. Radio propagation effects are incorporated by employing ray tracing models. Stamatelos et al. introduce a two-phase heuristic approach to solving the optimization design problem. In the first phase, given the number of APs required in the service area and a fixed power level, they determine a location of each AP. The AP placement is optimized with respect to the objective function and the initial power level. In the second phase, the optimal power level is determined while the AP locations are fixed with the value obtained from the first phase. However, the authors do not take into account user traffic, data rate density requirements, or capacity. They consider only a single floor (2-D) design. Medium access control is not considered. Access to the network depends only on the physical properties of the signal.

Rodrigues et al. [31] consider WLAN design for a service area covering two floors of a selected building. They aim to determine the best location and frequency channel for an AP. They divide the total area into small quadrangular pieces of demand areas. Each demand area is associated with a specific service priority that indicates how important it is to have good signal strength over that demand area. Candidate locations to install APs are selected based on two criteria: low installation cost and good attendance area. Six candidate locations are chosen. A path loss model is not employed to capture radio coverage. Instead, signal strength from each candidate AP to each demand points is measured. The authors address the problem with a two-phase design technique using two integer linear programming models. The first addresses the AP placement problem, while the second assigns frequency channels. In the first phase, given the number of available APs, the best location is determined by an objective function that maximizes overall signal strength in the service areas and considers the assigned service priorities. The second phase assigns frequency channels using an objective function that minimizes interference between APs by maximizing the sum of channel distances between adjacent APs where the minimum distance of three channels is required. However, the authors consider only signal coverage in the service area and fail to account for user demand and the traffic profile. AP locations are limited to a few opportunistic locations. The design is limited to a service area of two floors. The radio coverage characteristics are based on a limited number of measurements.

Unbehaun [32] extends the work of Anderson et al. [26] and Sherali et al. [28] by implementing the coverage planning for indoor WLANs as an optimization algorithm using a convex combination of two objective functions as proposed by Sherali et al. [28]. Unlike Sherali's solution technique [28], Unbehaun applies the simulated annealing approach used by Anderson et al. [26] to solve the design problem. Unbehaun separates the solution technique into two consecutive steps, propagation prediction and the coverage optimization, due to the complexity of the ray-tracing propagation model used in the paper. In the first step, he defines grid points of candidate locations for APs and computes the propagation data for all grid points. Then, he determines AP locations that optimize radio coverage by using a simulated annealing algorithm. The design in this paper considers three different frequency bands (5, 17, and 60 GHz). Two groups of experiments were conducted: single-transmitter and multiple-transmitter placement. The results of the single transmitter placement experiment show that in low frequency, 5 GHz, the optimal coverage planning is less sensitive to the AP location. This means that there appear to be several choices for AP locations that give close-to-optimum coverage. In the higher frequency (60 GHz), however, only a few locations yield good coverage. In other words, to achieve optimal or close-to-optimal coverage, the AP must be located at a few specific places. In the multiple-transmitter placement experiment, at 5 GHz, Unbehaun observed that the coverage optimization approach yielded only a small gain in signal coverage compared to the regular installation approach in which APs are located in the center of cells and placement is determined by dividing the service area into equal-sized rectangles [18]. At higher frequencies of 17 GHz and 60 GHz, however, the coverage optimization approach improved the worst-case signal quality by more than 12 dB while the average signal coverage was only slightly improved. Although this paper provides good insight into the efficiency of coverage optimization methods on the system in different frequency bands, it only takes into account the signal coverage objective and fails to consider user density and traffic demand in the service area.

2.2 Cellular Network Design

The design of cellular networks has usually followed the following procedure. Given the type of modulation scheme, the multiple-access technique and the required SIR for an acceptable service quality, the allowable transmission power and radio propagation characteristics are used to determine the coverage of a base station in a given area. The number of base stations required to be deployed to meet initial subscription demand is determined and approximate locations for the base stations are chosen.

In the past, the major design criterion of cellular networks was the coverage area. Conventional mobile engineering methods like the analytical approach [33, 34] are mainly focused on providing the best possible radio signal at every location of the planning region. The capacity aspects of network design are addressed only in later stages of the planning process. In this case, the radio frequency design objective and the network capacity objective are treated separately, so trade-offs between design objectives and optimal design are hard to obtain.

Now that mobile radio has made the transition into a mass-communication system, cost has become an important issue in system design. Since demand coverage can be viewed as revenue coverage, the network demand objective has become one of the major keys in cellular network design. To overcome the disadvantages of the conventional approach, the new design approach incorporates the demand criterion in the early stages of the design process [35].

Tutschku [35] proposed an integrated approach, in which the network design begins with the analysis of the expected teletraffic and this information is used together with other objectives (e.g., radio coverage and network deployment) in the radio network optimization. The author suggested a new way to represent the demand for teletraffic by a discrete point. The discretization of the traffic demand is called the Demand Node Concept [36, 37] where a demand node represents the center of an area that contains a quantum of demand (i.e., a fixed number of call requests per time unit). The demand nodes are generated by a partition clustering method, which uses the information in geographical information systems to estimate the distribution of teletraffic demand. Such a demand model incorporates factors such as land use, population density, and vehicular traffic. The partition clustering method [36] is based on the idea of dividing the service area until the teletraffic of every tessellation piece is below a threshold and the center of each tessellation piece represents a demand node. Thus, the demand nodes would be dense in areas of high traffic intensity (e.g., in the center of a city or on highways) and sparse in areas of low traffic intensity (e.g., in rural areas) The obtained spatial traffic distribution is then incorporated into the cellular network design via a mathematical formulation as a set covering problem, which aims to determine the number of base stations and their locations such that all demand points are served with an adequate signal strength [35].

2.3 Demand Node Representation for Traffic Demand in Cellular Network Designs

The idea of a spatial pattern of demand has been used widely in facility location problems [38]. It provides information on the geographic pattern of demand for retail goods and services. Typically, the demand within a small geographic area is estimated from population, income, and demographic characteristics and assigned to a point within the area.

The concept of a demand node has been applied in the cellular network designs to impose network capacity requirements in the design processes [35, 39]. The expected traffic demand in the service area is estimated based on the geographic and demographic information of the service areas.

There are two types of demand node representations that have been used in the cellular network designs in existing studies: the uniform size demand node and the variable size demand node. These demand node representations are explained in more detail below.

2.3.1 Uniform-Size Demand Node

For the uniform-size demand node, each node represents the center of an area that contains the same amount of traffic demand (i.e., a fixed number of call requests per time unit) [36]. The demand nodes are dense in areas of high traffic intensity and sparse in areas of low traffic intensity. Figure 2.3 illustrates the uniform-size demand node representation as it is applied in the cellular network design presented by Tutschku [35].

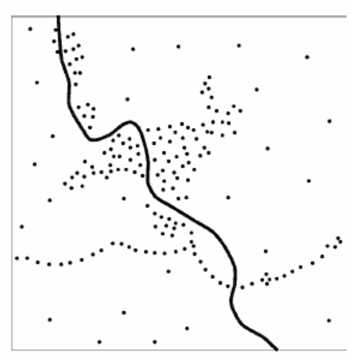


Figure 2.3 Uniform size demand node [36]

2.3.2 Variable-Size Demand Node

A demand node in this model has a fixed location and represents a certain number of calls per time unit. However, this number can vary across demand nodes. The demand nodes are large in areas of high traffic intensity and small in areas of low traffic intensity. Figure 2.4 illustrates the demand node representation of this type which is applied in the base station transmitter placement presented by Weicker et al [39].

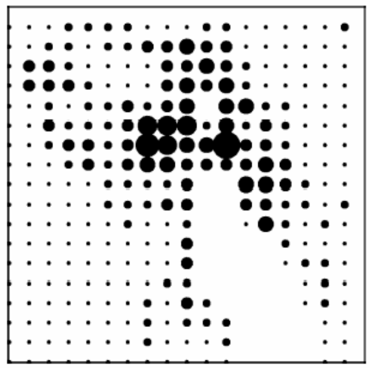


Figure 2.4 Variable-size demand node [39]

2.4 Network Usage Characteristic in WLANs

Network usage characteristics in WLANs have been studied widely in various environments. Tang and Baker [40] observed WLAN usage characteristics in a university building. Kotz and Essien [41, 42] and Hutchins and Zegura [43] investigated network usage of a university WLAN ranging across multiple campus buildings. Balachandran et al. [44] examined usage of a WLAN in a large auditorium that hosted conferences. Balazinska and Castro [45] conducted network traces on a large corporate environment ranging across three buildings. From the existing network trace studies, WLAN usage characteristics are summarized as follows:

- **User Mobility and User Population**

WLAN users are rather stationary, spending most of the time they are connected to the network at a single location [40, 45]. Wireless users in corporate and academic offices usually perform data transfer activities at their office desks, occasionally moving to conference or meeting rooms. Researchers have reported that the number of wireless users in a network varies according to daily and hourly work, meeting, and/or class schedules [40, 44, 45]. In a corporate network environment, the number of user follows an office work hour pattern, and there is a reduction in the number of users around lunch time [45]. This network usage pattern is similar to what is observed at administrative and office locations in the university environment [40, 41]. In auditoriums or conference rooms, the number of users of APs varies with the schedule of activities [44].

Thus, the number of wireless users handled by APs varies based on the APs' locations. In a place where wireless users usually congregate (e.g., auditoriums, classrooms or cafeteria), the APs see a high number of users connecting to the network [40, 41, 45]. In private offices and doom rooms, it appears fewer users connect to the APs [41, 42].

- **Correlation between Number of Users and Traffic Load**

Researchers report that little correlation exists between the number of users and the traffic over APs. Balazinska and Castro [45] and Balachandran et al. [44] observe that in locations such as auditoriums and cafeterias where many people congregate and connect to APs, most users are passive for a large part of their sessions, occasionally checking email. Kotz and Essien [41, 42] also notice little correlation between the number of users and the amount of traffic going through APs. They find that the majority of network users are located in classrooms while the majority of network traffic is generated by those in dorm rooms and graduate offices.

- **Correlation between Locations and Traffic Load**

Researchers report a correlation between locations where users are present (and, implicitly, the locations of APs) and traffic load over APs. Users pursue different types of activities, in part based on their location, and this in turn affects their data transfer activity. Balazinska et al. and Balachandran et al. [44, 45] found that location significantly affects users' level of data transfer activities. They observed that users accessing the network while in auditoriums have low data transfer activities, only occasionally check e-mails. Other network trace studies also find a relationship between location and data transfer activity level. Balachandran et al. [44] found that while approximately 50% of wireless users connect to APs during conference talks, 80% of those who connect are inactive for more than 30% of their session length. Kotz and Essien [41, 42] find more data transfer activity from graduate offices and doom rooms than from other more crowed locations (e.g., classrooms, cafeterias and libraries).

2.5 Quality of Service in WLANs

As the use of WLANs extends beyond simple data transfer to multimedia applications, the need to address Quality of Service (QoS) issues becomes critical. QoS in WLANs has been studied with regard to timeliness (e.g., delay, jitter) and bandwidth (e.g., system data rate, application level data rate) [46]. Several recommendations have been made for improvement in the timeliness of the original 802.11 MAC protocol. Most of the work in this area has focused on analyzing performance [11], tuning the 802.11 MAC parameters [47], and introducing a priority queuing scheme [48]. To address the bandwidth QoS, modulation and coding techniques such as Orthogonal Frequency Division Multiplexing (OFDM) [49], Complimentary Code Keying (CCK) [50], and Packet Binary Convolutional Code (PBCC) [51] have been developed to improve system data rates and frequency spectrum utilization. However, despite the enhanced MAC protocols and advanced modulation and coding techniques, WLANs might not be able to guarantee that the QoS that users' applications require is met if the WLAN system is overloaded with large numbers of users try to access the same AP [13].

2.6 Constraint Satisfaction Problem

A Constraint Satisfaction Problem (CSP) consists of a set of variables (***V***), a set of domains associating to the variables (***D***) (i.e., a set of all possible values that can be assigned to the variable), and a set of constraints (***C***). The set of constraints puts restrictions on the values that variables in the set ***V*** can take simultaneously. A feasible solution to a CSP is an assignment of values from associating domains to all the variables such that no constraint is violated.

Tsang [52] classifies CSPs into four categories based on an application's requirements: (a) CSPs in which one has to find any feasible solution, (b) CSPs in which one has to find all feasible solutions, (c) CSPs in which one has to find an optimal solution, and (d) CSPs in which the constraints are so tight that it is difficult or impossible to satisfy all of them and one has to find a solution that satisfies as many constraints as possible. In this dissertation, we focus on the task of finding any feasible network configuration solution to the demand-based WLAN design problem.

There are two types of approaches to solving the CSP: systematic approaches and repair approaches, also known as iterative improvement approaches.

The systematic approaches typically develop a search tree based on the possible values for each of the variables of the CSP. Such search algorithms start from an empty variable assignment and enlarge the assignment step by step until all variables are assigned feasible values [52]. When a dead-end is reached, backtracking occurs. The primary limitation of this approach is that it can only handle small problems and is not large enough to accommodate practical problems [53]. This type of approaches is best for CSPs that do not need a complete assignment of variables to evaluate the constraints. For example, in scheduling problems the constraints can be evaluated after partially assigning values to some variables because the constraints may only involve a single restriction of each variable.

In repair approaches, an initial solution is assigned to all variables regardless of feasibility. This solution is gradually repaired in order to reduce the infeasibility until all constraints are satisfied. The repair approach incorporates variations of hill climbing or local search techniques such as tabu search [53] and genetic algorithms [54]. The repair approaches are best for the CSPs that require a complete assignment of all variables so that the constraints can be evaluated. For example, in wireless network design problems, one can only evaluate the signal quality requirements in the service areas after assigning values to all transmitters' parameters. The solution technique developed in this dissertation pursues the repair type approach and uses tabu search as a basic mechanism to carry out the repairing process.

2.7 Path Loss Models for Indoor Radio Propagation

Indoor radio propagation is an important issue for WLAN system design. The coverage area of an AP can be estimated by using path loss models, which consider the impact of environment factors such as building layout, construction materials, and obstructions in predicting indoor propagation [12]. Rappaport [12] provides surveys on indoor path loss models and classifies the existing models into two categories: single floor and multi-floor models.

2.7.1 Single Floor Models

Single floor models are used when the transmitter and receiver are located on the same floor. Examples of single floor models that can be found in the literature are as follows:

Log-Distance Model:

The log-distance model [12] considers that path loss increases with distance from the transmitter. The path loss model is expressed as:

$$PL(d) = PL(d_0) + 10n\,log_{10}\left(\frac{d}{d_0}\right) + X_\delta \tag{2.1}$$

where d is the distance from the transmitter, $PL(d_0)$ is the path loss at reference distance d_0 , n is the path loss exponent specifying the path loss behavior for a particular environment, and X_δ is a log-normally distributed random variable representing the shadow fading with standard deviation δ dB.

Partition-Dependent Model:

The partition-dependent model [29] considers an explicit number of partitions existing between the transmitter and receiver. It assumes that signal attenuation is in the free space (n =2) plus additional path loss increasing with the number of partitions. The path loss model is expressed as:

$$PL(d) = PL(d_0) + 10n\,log_{10}\left(\frac{d}{d_0}\right) + \sum m_i w_i + X_\delta \tag{2.2}$$

where m_i is the number of partitions of type i and w_i is the attenuation factor in dB for a partition of type i.

2.7.2 Multi-Floor Models

Multifloor models are used only when the locations of the transmitter and receiver are located on different floors of a building. Motley and Keenan [55], Seidel and Rappaport [29] proposed similar floor-dependent path loss models. A general formula for computing the mean path loss in a multi-floor scenario is expressed as:

$$\overline{PL}(d) = PL(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) + FAC \tag{2.3}$$

where *FAC* is the floor attenuation component, a function of the number of intervening floors. In the Motley and Keenan model, the *FAC* (dB) increases linearly with the number of floors (i.e., each floor reduces the signal strength by the same amount). However, Seidel and Rappaport [29] have observed that different number of intervening floors attenuate the signal strength by different amounts. Their *FAC* increases non-linearly with the number of floors. For example, *FAC*s of 12.9, 18.7, 24.4, and 27 dB are corresponding to signal propagation through 1, 2, 3 and 4 intervening floors, respectively [29]. Andersen and Rappaport et al. [56] report no significant increase in *FAC* for more than five floors of separation.

2.8 Antenna Patterns

The IEEE standard 145-1993 defines an antenna pattern as "a mathematical function or graphical representation of the radiation properties of the antenna as a function of space coordinates." Power pattern, one method of measuring radiation properties of the antenna, is a trace of the measured power at a constant distance from a transmitting antenna. Typically, antenna performance is measured in terms of gain in dBi, denoting power measured in reference to an isotropic radiator. Since the isotropic radiator is assumed to have unit gain, the term dBi is often used interchangeably with dB.

Antenna patterns are represented using the spherical coordinate system shown in Figure 2.5. The horizontal (x-y) plane, called the azimuth plane, is denoted by $\theta = 90°$. The vertical planes are called elevation planes. Two elevation planes of particular interest are $\phi = 0°$ (x-z plane) and $\phi = 90°$ (y-z plane). The relationship between rectangular and spherical coordinates is:

$x = r \sin\theta\cos\phi$, $y = r \sin\theta\sin\phi$, $z = r \cos\theta$, where $r \geq 0$, $0° \leq \phi < 360°$, $0° \leq \theta \leq 180°$,

and $\quad r = \sqrt{x^2 + y^2 + z^2}$, $\quad \theta = tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right)$, $\quad \phi = tan^{-1}\frac{y}{x}$.

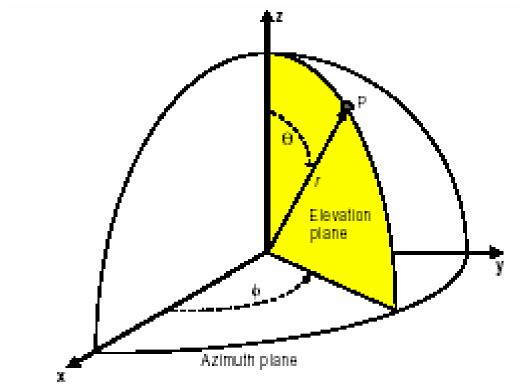


Figure 2.5 Spherical coordinate system [57]

(Note that the point P may be represented using Cartesian coordinates *(x, y, z)* or by using spherical coordinates *(r, θ, ϕ)*.)

2.8.1 Antenna Gain Equations

Antenna gain, $G_T(\theta, \phi)$, as seen by a receiving point P (as shown in Figure 2.5) is expressed by the following equation (2.5) [57].

$$G_T(\theta,\phi)[in\ dB] = G_B + 20\log_{10}|G_M(\theta,\phi)| \quad (2.5)$$

where

- G_B is azimuth gain or the peak directivity in dB.
- $G_M(\theta, \phi)$ is the normalized electric field in the elevation plane.

2.8.2 Antenna Patterns

Antenna patterns can be classified into three types:

1) Isotropic Pattern

An isotropic antenna has equal radiation, and the gain equation $G_T(\theta, \phi)$ is constant in all directions [58]. In three dimensions, the radiation pattern is spherical in shape. Such an ideal radiator is often used as reference to express the directional properties of other antennas. Figure 2.6 illustrates the concept of directivity, which is a figure-of-merit to describe the directional properties of an antenna. The direction of peak directivity (i.e., maximum radiation from the antenna) is called the boresight direction.

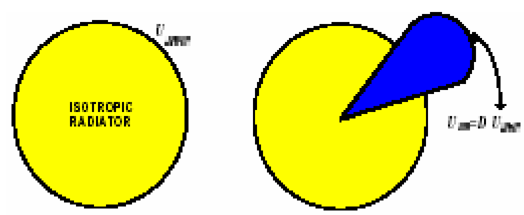


Figure 2.6 Illustration of directivity.
Note that both circles represent the average radiation intensity of an isotropic source. The darker region represents the directional property of the antenna.

2) Omnidirectional Patterns

An omnidirectional pattern has a donut shape in three dimensions [59]. Antennas with these patterns are the most widely used for indoor communications. The pattern is circular and constant in the azimuth (x-y) plane and directional in any elevation plane. The generalized gain equation is written as [59]:

$$G_T(\theta,\phi)[in\ dB] = G_{AZ}(\phi)[in\ dB] + G_{EL}(\theta)[in\ dB] \tag{2.6}$$

where $G_{AZ}(\phi)$ is a constant and $G_{AZ}(\phi) = G_B$. $G_{EL}(\theta) = G_M(\theta, \phi)$. Figure 2.7 is a simple illustration of the pattern with the donut sliced to show an elevation plane pattern.

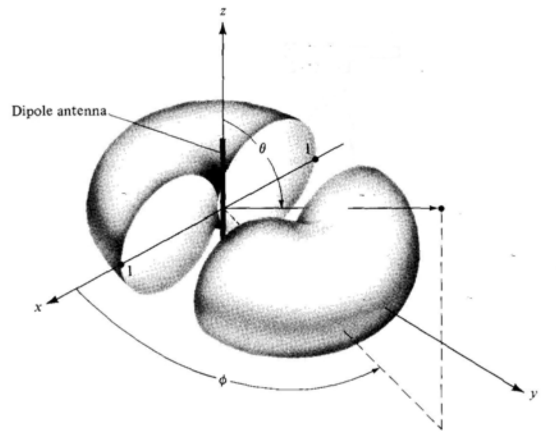


Figure 2.7 Omnidirectioanal antenna pattern [59]

Types of omnidirectional antennas that are commonly used in 802.11 WLANs include dipole antennas and monopole antennas [60].

2.1) Dipole Antennas

Types of dipole antennas include the halfwave dipole and quarterwave dipole antennas.

2.1.1) Halfwave Dipole

The halfwave dipole antenna has the physical length of a resonant wire of $\lambda/2$, where λ is the wavelength. The halfpower beamwidth of the halfwave dipole is 78°. The peak directivity of the halfwave dipole is in the azimuth plane, and it is constant in all directions around the plane. Figure 2.8a shows the pattern of the elevation plane ($G_{EL}(\theta)$) of the halfwave dipole antenna. The antenna patterns can be written using the following closed-form equations [57]:

Azimuth plane: $G_{AZ}(\phi)$ = peak directivity, constant (in dB) (2.7)

Elevation plane: $$G_{EL}(\theta) = 20 log_{10} \left| \frac{cos(0.5\pi\, cos\,\theta)}{sin\,\theta} \right| \text{ dB for all } \theta \in (0°, 180°) \tag{2.8}$$

2.1.2) Quarterwave Dipole

The quarterwave dipole has the physical length of a resonant wire of λ/4. The pattern is similar to the halfwave dipole except that the halfwave beamwidth of a quarterwave dipole is 87°. The peak directivity is constant in the azimuth plane. Figure 2.8b depicts the elevation plane pattern of the quarterwave dipole. The antenna patterns can be written using the following closed-form equations [57]:

Azimuth plane: $G_{AZ}(\phi)$ = peak directivity and constant (in dB) (2.9)

Elevation plane: $$G_{EL}(\theta) = 20 log_{10} \left| \frac{cos(0.25\pi\, cos\,\theta) - cos(0.25\pi)}{sin\,\theta} \right| \text{ dB} \tag{2.10}$$

for all $\theta \in (0°, 180°)$

a. Halfwave dipole b. Quarterwave dipole

Figure 2.8 Antenna pattern of the elevation plane of the dipole omnidirectional antenna [57]

2.2) Monopole Antennas

The monopole is a dipole divided in half at its center feed point and fed against a ground plane, as shown in Figure 2.9. However, high frequency monopoles are usually fed from behind the ground plane, as shown in Figure 2.9. A monopole above the ground plane radiates one-half the total power of a similar dipole in free space. Thus, the directivity of a monopole is double that of a similar size dipole.

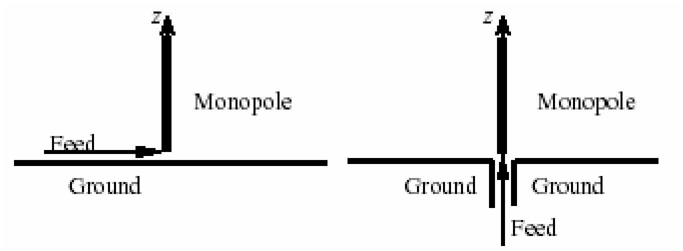


Figure 2.9 Monopole antennas [57]
(The figure depicts two ways in which energy can be fed into the antenna.)

The three-dimensional form of the monopole antenna is a half donut shape where the donut is chopped along the azimuth plane. In this case, it is assumed that the ground is both infinite and a perfect electrical conductor. Elliott [61] and Balanis [59] base their closed-form monopole equations on the dipole equations, modified to fit ground plane effects of a monopole. They multiply the closed-form equations for the dipole by the factor "k sin(n sinθ"), where n and k are determined by the size of the ground plane and the size of the monopole antenna [59]. This factor is called the "ground plane correction factor."

2.2.1) Halfwave Monopole

The peak directivity is constant in the azimuth plane, and it is denoted by the antenna specification. The equation for the elevation plane is modified from that of the fullwave dipole by multiplying with the "ground plane correction factor". The elevation plane plot is shown in Figure 2.10a. The closed form equation of the halfwave monopole is expressed as [57]:

Azimuth plane: $G_{AZ}(\phi)$ = constant (in dB) (2.11)

Elevation plane: $$G_{EL}(\theta) = 20\,log_{10}\left|\frac{cos(\pi\cos\theta+1)}{sin\theta}\right|\left|k\,sin(n\,sin\theta)\right| \text{ dB} \tag{2.12}$$

for all $\theta \in (0°, 90°)$ with n = 2.75 [57] and k = 1.2015 [57]
Elsewhere, $G_{EL}(\theta)$ = -60 dB [57]

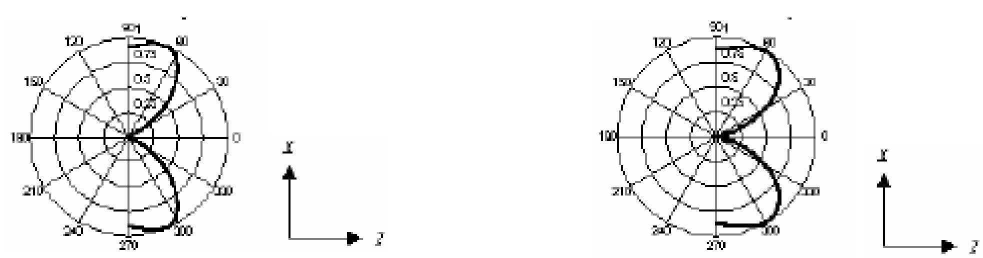
2.2.2) Quarterwave Monopole

As in the halfwave monopole, the peak directivity is constant in the azimuth plane, and is denoted by the antenna specification. The elevation plane is modified from the halfwave dipole by multiplying with the "ground plane correction factor". The elevation plane plot is shown in Figure 2.10b. The closed form equation of the quarterwave monopole is expressed as [57]:

Azimuth plane: $G_{AZ}(\phi)$ = constant (in dB) (2.13)

Elevation plane: $$G_{EL}(\theta) = 20\,log_{10}\left|\frac{cos(0.5\pi\cos\theta)}{sin\theta}\right|\left|k\,sin(n\,sin\theta)\right| \text{ dB} \tag{2.14}$$

for all $\theta \in (0°, 90°)$ with n = 2.5 [57] and k = 1.4782 [57]
Elsewhere, $G_{EL}(\theta)$ = -60 dB [57]

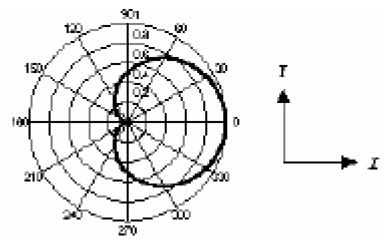


a. Halfwave monopole b. Quarterwave monopole

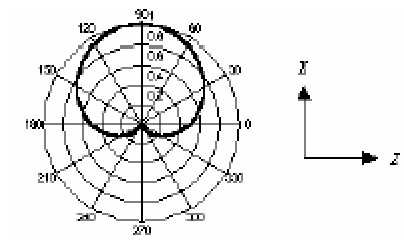
Figure 2.10 Antenna pattern of the elevation plane of the monopole omnidirectional antenna [57]

3) Directional Antenna

A directional antenna radiates and receives electromagnetic waves more efficiently in certain directions than others. Patch antennas are typically used in indoor 802.11 WLANs [60]. Figure 2.11 shows an example of the patch antenna pattern.



a. Azimuth plane



b. Elevation plane

Figure 2.11 Antenna pattern of the patch directional antenna [57]

2.9 Summary

This chapter presents a literature review of indoor wireless network designs, particular WLAN design, and cellular network designs. The chapter also provides an overview of related issues including network usage characteristics, quality of service in WLANs, constraint satisfaction problems, path loss models and 3D antenna patterns. A review of the current literature shows several shortcomings and missing components of existing WLAN design approaches. Most of these existing methods are based on trial and error and measurement-based approaches. These methods are time consuming and labor intensive. Existing optimization based approaches limit their focus to optimizing signal strength in the service area. The present design techniques do not consider user density, traffic demand, or network capacity requirements in the design process. Additionally, most of the existing WLAN design approaches limit their focuses to network design for small and single floor service areas. A systematic design for multi-floor service scenarios is needed. This need is addressed in this dissertation.

3.0 DEMAND-BASED WLAN DESIGN: MODEL AND PROBLEM FORMULATION

Wireless local area networks (WLANs) are experiencing tremendous growth, providing untethered data networking capabilities [2, 4]. Their deployment has been facilitated by the availability of unlicensed frequency spectrum and inexpensive network equipment [1]. As WLAN-access devices become cheaper, smaller and more powerful, there has been phenomenal growth in the number of people who use WLAN services [1-4]. In designing WLANs, network designers must accommodate both growth of the user population and increased demand for services. Unfortunately, most existing indoor wireless network design methods limit their focus to radio signal coverage, which only ensures that an adequate signal strength is maintained in the intended service area [14-24, 26, 28, 30-32]. In this dissertation, we present a demand-based WLAN design approach, incorporating traffic demand issues into the network design model.

This chapter describes in detail a demand-based WLAN design approach. Section 3.1 gives the problem definition. Section 3.2 presents network design criteria imposed in the proposed model. Section 3.3 discusses a demand node representation for WLAN design. Section 3.4 presents the mathematical problem formulation of the demand-based WLAN design model. Section 3.5 provides an overview of the heuristic solution technique developed to solve the demand-based WLAN design problem. Finally, section 3.6 summarizes the chapter.

3.1 Problem Definition

The task of demand-based WLAN design is to place a sufficient number of access points (APs) in a service area that may be located on a single floor or range across multiple floors. The APs may be configured with different power levels and frequency channels. The power level and frequency channel of an AP, together with the environment specific path loss and an antenna radiation pattern, determines the region (called Basic Service Area (BSA)) in which the AP can support traffic demand to/from wireless users. According to capacity analysis of the CSMA/CA protocol used in WLANs, the capacity of an AP can vary depending on the number of wireless users simultaneously transferring data through the AP. As the number of wireless users with active data transfer connections to a particular AP increases, the effective AP capacity decreases. Thus, the number of APs in a service area should be a function of the number and density characteristics of the wireless users. Due to the low cost of the APs compared to that of the wireless devices with which they are communicating, minimizing the number of the APs -- as would be recommended by current optimization approaches -- is unnecessary [13]. Thus, we formulate the WLAN design problem as a Constraint Satisfaction Problem (CSP) which aims to determine the number of APs, identify their locations, and assign power levels and frequency channels to them such that the resulting network satisfies the network design criteria described below.

3.2 Network Design Criteria

3.2.1 Radio Signal Coverage Requirement

In wireless network design, a fundamental requirement for a network is that it provide radio signal coverage over a target service area [62]. Wireless users located in the service region require a certain level of radio signal quality in order to be able to access the network. As a measure of radio signal availability and coverage, the received signal strength and the signal to interference ratio (SIR) are considered in the design model. The received signal strength by wireless nodes must exceed the specified receiver sensitivity threshold. Additionally, the received signal strength from the serving AP must be sufficiently greater than the signal received from other APs operating on the same or overlapping frequency channels as specified by the SIR.

3.2.2 User Data Rate Capacity Requirement

An obtainable user data rate (link rate) has become an essential concern in network planning for WLANs as the user population grows and multimedia applications requiring higher data rates spread [13, 14]. We incorporate a mean user data rate requirement in the developed demand-based WLAN design model.

As discussed in section 2.4, average available user data rates observed in the network trace studies does not depend just on the number of wireless users existing in the service area, but also on the activity of wireless users in the network [40-42, 44, 45].

There are correlations between user behaviors and traffic volume in the network [45]. The user behavior in turn correlates to types of locations where users are situated and the major activities users typically pursue in such locations [40-42, 44, 45]. We incorporate this critical information about characteristics of WLAN usage and traffic patterns in our design model as described in the following sections.

3.3 Demand Node Representation for WLAN Design

The demand node concept has been used widely in facility location problems to provide information on the geographic pattern of demand for retail goods and services [38]. This concept has recently been applied in wide-area wireless network designs to represent the distribution of expected network traffic in a service area [35, 39]. In wide area wireless networks, demand nodes are derived from land usage and demographic information. As described in section 2.3, two types of demand nodes have been applied in cellular network design: a uniform size demand node and a variable size demand node. The uniform size demand node represents the center of an area that contains a fixed number of call requests per time unit while the variable size demand node represents a certain number of calls per time unit in which the number can vary across demand nodes. The cellular network designs aim to provide sufficient traffic channels to accommodate the estimated number of call requests represented by demand nodes.

Here we apply the demand node concept to WLAN design. However, in the context of WLAN a demand node is different from that used in the cellular network design as defined in the following discussion.

3.3.1 Definition of a Demand Node for WLAN Design

In WLAN design, a demand node represents an individual prospective wireless user in a service area. In local area network design, average locations of prospective wireless users can be estimated simply from available seats and working desks provided in service areas. The facility administrators can provide information about the estimated number of prospective wireless users in the area.

The reasoning for adopting a different definition of a demand node in WLAN design is that more precise information about the potential number of wireless users and users' locations is needed in order to appropriately place APs and assign users to APs. This is because in WLANs, users communicate through APs using the CSMA/CA protocol in which users compete for channel access and share AP capacity. In this case, information about the number of users is required to calculate the potential data rate and average AP capacity while information about user locations is needed to appropriately assign users to an AP based on an acceptable radio signal level.

Another consideration in wireless network design is traffic fluctuation with the time of day [63]. As mentioned in section 2.4, researchers have observed that the volume of network traffic and distribution of users changes with daily class and work schedule as users move between classrooms or move from lab facilities to offices [40-42, 44, 45]. As an example, Figure 3.1 shows the distribution of prospective wireless users during different times of day on the fourth floor of the School of Information Science (SIS) building at the University of Pittsburgh. Between 9AM and 12PM, students are present in classrooms 403, 406, 409, and 411. At other times, group of students are present in different rooms. In our design approach, we take into account the variation of user density and distribution by applying a worst-case (peak) estimate approach [64] in which a representation of overall user distribution in a service area is drawn from the user population when the highest user density may be present in each sub-area or room. This approach constitutes a busy-hour user population representation, capturing user density fluctuations due to mobile user time of day characteristics. Figure 3.2 illustrates the resulting user population representation of the fourth floor SIS building. The obtained busy-hour user population information will be used in the network design process.

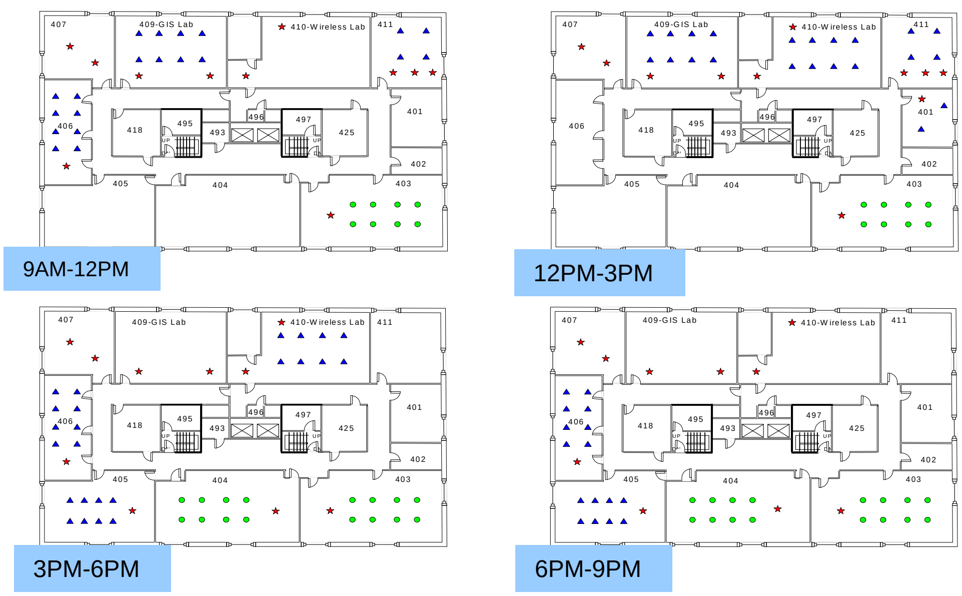


Figure 3.1 User distribution during different times of day

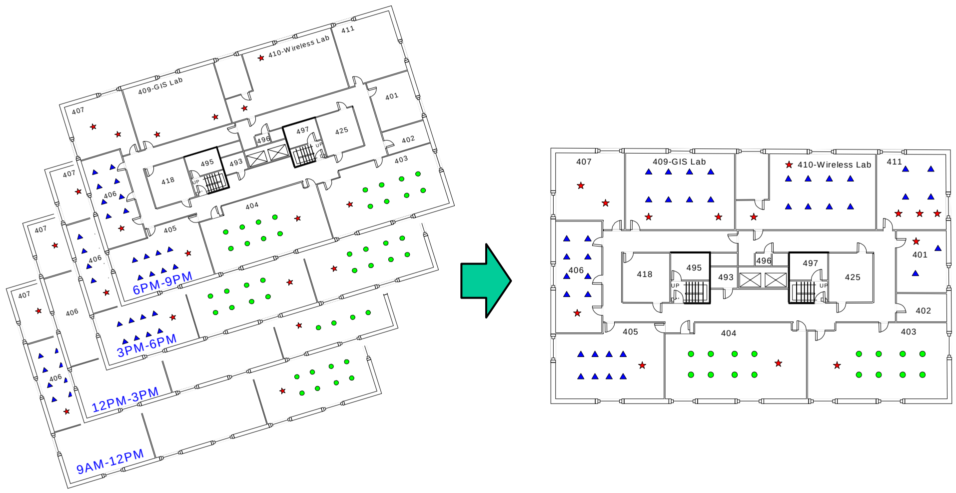


Figure 3.2 User population representation for WLAN design

While estimates of demand node locations representing prospective wireless users in service areas can be derived and represented by the busy-hour user population as described above, expected traffic volume of each demand node on the network can be estimated from network usage characteristics observed in the existing WLAN.

3.3.2 WLAN Usage Characteristics and User Activity Level

Several network trace studies characterize the usage of WLANs in various environments such as on university campuses [41, 42], in corporate office buildings [45], in academic buildings [40], and in a large auditorium [44]. Researchers have found similarities in network usage characteristics among different network environments [40-42, 44, 45]. It has been observed that traffic load at APs does not depend entirely on the number of wireless users who situate within radio coverage of APs and associate to them. The amount of traffic depends on users' level of data transfer activity on the network as well. Network trace studies reported in the literature show a correlation between users' level of data transfer activity and locations where users are present [41, 42, 44, 45].

As an example, network traces studies at Dartmouth College [41, 42] show that the highest number of wireless users associated with APs at any time occurs in the areas where users usually congregate, such as in classrooms and auditoriums. However, the highest amount of traffic does not occur in these areas. Rather, the places that experience high traffic volume are private offices and dorm rooms where a smaller number of wireless users are present [41, 42]. Presumably, most people attending classes, meetings, or conferences usually concentrate on the event, only occasionally checking e-mail and/or downloading slides. Balachandran et al. [44] report that wireless users attending meeting or conference talks tend to be idle over large portion of their session duration. They found that only about 50% of wireless users connect to APs while attending meetings or conference talks and about 80% of those who connect are inactive more than 30% of their session length. The average user data rate observed in such an environment tends to be low (between 15 and 80 Kbps) [44].

To account for the correlation between network usage characteristics and locations where users situate in our demand-based network design model, we categorize sub-areas of the whole service area into two types. These types are based on typical tasks corresponding to particular locations and network usage characteristics of the sub-areas. They are the private sub-area and the public sub-area.

- Private sub-area:

Private sub-areas are sub-areas of a service area that are limited to authorized persons only. Examples of private sub-areas include office areas, homes, and dorm rooms. In these areas, the number of wireless users and their locations can be known or expected a priori.

- Public sub-area:

Public sub-areas are locations where large numbers of users often congregate. The number of wireless users varies depending upon events or schedules. We subdivide the public sub-areas into two categories: schedule-based activities and unschedule-based activities. The public sub-areas for schedule-based activities exhibit a timeline during which the group of wireless users exists for a specific purpose. Examples of these locations include classrooms, meeting rooms, auditoriums, and cafeterias. The public sub-areas for unschedule-based activities have a poorly structure timeline; users occasionally enter and leave. These locations include libraries, student lounges, and coffee houses.

We account for the correlation between network usage characteristics and locations where users situate in the demand-based WLAN design model through a parameter called "user activity level" which is defined as follows:

User activity level (α_t) is the percentage of wireless users in a sub-area of type t who are simultaneously transferring data over APs. These active users are participating in medium contention to gain access to a communication channel and share AP capacity. We consider the remaining users ($1-\alpha_t$) to be idle users who are situated in a sub-area of type t but do not generate data transfer activity over the network at a particular time. The idle users do not affect AP capacity [44].

The busy-hour user population and user activity level (α_t) parameter are incorporated into the demand-based WLAN design model formulated in the next section.

3.4 Demand-based WLAN Design Model

Here we develop a new model for the WLAN design problem that incorporates not only radio signal coverage requirements but also network capacity requirements. We mathematically formulate a demand-based WLAN design model as a Constraint Satisfaction Problem (CSP). First we define terms and notations used in the problem formulation.

3.4.1 Definition

3.4.1.1 Signal Test Points

The service area where the WLAN is to be deployed is divided into a finite number of grid points. Those points within the usage area are referred to as signal test points (STPs) where the received signal strength and the SIR level will be assessed. Note that grid points in places which are defined as non-usage areas (e.g., elevator shafts and restrooms) are not included in the set of STPs.

Let $G = \{g_1, g_2, \ldots, g_c\}$ denote a set of STPs representing locations where the received signal strength and the SIR level will be tested, and c is the total number of

STPs in the service area. Each STP g_h refers to a coordinate in three-dimensional space (x_h, y_h, z_h), where $1 \leq h \leq c$ and z_h is the floor where g_h is located.

3.4.1.2 Demand Nodes

A set of demand nodes represents prospective wireless users that may be present in the service area. It represents a busy-hour user population in service areas.

Let $U = \{d_1^t, d_2^t, \ldots, d_m^t\}$ denote a set of demand nodes, given by their position inside the service area, where m is the total number of demand nodes in the set U. Index t indicates the type of sub-area where demand node i is located: $1 \leq i \leq m$ and $t \in T$ where T is the set of sub-area types. Here we define three types of sub-areas: $T = \{1,2,3\}$ where 1 denotes private sub-areas, 2 denotes public sub-areas for unscheduled activities, and 3 denotes public sub-areas for schedule-based activities. The position of demand node i within the service area is denoted by (x_i, y_i, z_i), where (x_i, y_i) is the coordinates on floor z_i where d_i^t is located.

We incorporate network usage characteristics of demand nodes located in sub-areas of type t through the user activity level (α_t) and the average data rate requirement (R_t). The set of demand nodes together with the sub-area classification and parameters specifying network usage characteristics (α_t and R_t) are given as input to the design process by the designer.

3.4.1.3 Network Configuration

The demand-based WLAN design problem aims to determine a network configuration such that the radio signal is available throughout the defined service area (represented by a set of STPs, G) and the network capacity is adequate to accommodate expected traffic demand from a set of demand nodes, U. A network configuration specifies the number of access points (APs) and their parameters including locations, frequency channels, and power levels.

Let $A = \{ap_1, ap_2, \ldots, ap_n\}$ denote a set of APs used in the service area, where n is the total number of APs required. Let $ac_j = \{p_j, f_j, (x_j, y_j, z_j)\}$ denote a set of parameters assigned to ap_j for $1 \leq j \leq n$, where p_j denotes the power level assigned to ap_j, f_j denotes the frequency channel assigned to ap_j, and (x_j, y_j, z_j) denotes the coordinate (x_j, y_j) on floor z_j where ap_j is placed.

3.4.2 Notation

The following notation is defined and used in the mathematical formulation of the WLAN design problem:

(x, y, z)	=	Coordinate in three-dimensional space, where (x, y) represents a location on floor z
A	=	Set of APs used in the service area
U	=	Set of demand nodes in the service area
G	=	Set of signal test points (STPs) that are locations for testing received signal strength and SIR level
T	=	Set of sub-area types

Indices:

i	=	$1, 2, \ldots, m$ demand nodes in set U
j	=	$1, 2, \ldots, n$ APs in set A
h	=	$1, 2, \ldots, c$ STPs in set G

t = 1, 2, …, s sub-area types in set T

Variables:

n = Number of APs used in the network

p_j = power level of $ap_j \in A$

f_j = frequency channel of $ap_j \in A$

d_{ij}^{t} = User association binary variable that equals 1 if demand node $i \in U$ associates to $ap_j \in A$; 0 otherwise. Index $t \in T$ represents the type of sub-area where demand node i is located

g_{hj} = Signal availability binary variable that equals 1 if STP $h \in G$ can receive a signal from $ap_j \in A$; 0 otherwise

(x_j, y_j, z_j) = location of $ap_j \in A$, (x_j, y_j) represents the coordinates on floor z_j

Domains:

D_n = Integer number

D_p = Set of candidate power levels for variable p_j = {P_1, P_2,…, P_{max}}

D_f = Set of candidate frequency channels for variable f_j = {F_1, F_2, …, F_k}

D_d = {0, 1} the domain of binary variable d_{ij}^{t}

D_g = {0, 1} the domain of binary variable g_{hj}

$D_{(x,y,z)}$ = The domain of variable (x_j, y_j, z_j) where $x_{min} < x_j < x_{max}$, $y_{min} < y_j < y_{max}$, and $z_j \in FLOOR = \{1^{st}, 2^{nd}, \ldots, r^{th}\}$

3.4.3 Static Parameters

Static parameters depend solely on standard requirements and characteristics of the user activities in WLAN service areas and do not change during the design process. Static parameters include those specifying physical signal requirements (e.g., the received signal strength ($P_{Rthreshold}$) and the SIR level ($SIR_{threshold}$)), user profiles (e.g., the user activity level (α_t) and the average user data rate requirement (R_t)), adjacent channel interference between signals from overlapping channels (η_{jk}), and the data rate capacity of AP (C_j) that will be employed in the network.

Table 3.1 Static parameters

Parameter	Definition	Unit
α_t	User activity level defines the percentage of wireless users in sub-area type t that are engaged in a data transfer activities (i.e., participating in channel contention and sharing AP capacity)	-
R_t	Average user data rate requirement in the sub-area type t	bps
$P_{Rthreshold}$	Received sensitivity threshold	dBm
$SIR_{threshold}$	Signal to interference ratio threshold	dB
η_{jk}	Adjacent channel interference, specifying the percentage of interfering power that the signal operating at f_k dissipates to the signal operating at f_j	-
C_j	Data rate capacity of the ap_j	bps

3.4.4 Dynamic Parameters

Values of dynamic parameters are recomputed when the variables are set with different values during the course of the design process. In our design model, the dynamic parameters include received signal strength ($P_{R_{ij}}$), interference level ($Intf_{ij}$), and average obtainable data rate (r_i^t).

Table 3.2 Dynamic parameters

Parameter	Definition	Unit
$P_{R_{ij}}$	Received signal strength that user i receives from ap_j	dBm
$Intf_{ij}$	Interference level at user i, when associating to ap_j	dBm
r_i^t	Average data rate that user i who situates in sub-area type t can obtain	bps

- **Received signal strength ($P_{R_{ij}}$)**

The received signal strength ($P_{R_{ij}}$) that user i receives from ap_j, can be computed from:

$$P_{R_{ij}}(dBm) = p_j - L\big(f_j, (x_i, y_i, z_i), (x_j, y_j, z_j)\big) + G_A\big((x_i, y_i, z_i), (x_j, y_j, z_j)\big) \tag{3.1}$$

where

$L(f_j, (x_i, y_i, z_i), (x_j, y_j, z_j))$ is the path loss function between the location of ap_j, (x_j, y_j, z_j) and the location of user i, (x_i, y_i, z_i) when ap_j uses frequency channel f_j.

$G_A((x_i, y_i, z_i), (x_j, y_j, z_j))$ is the antenna gain based on the 3D antenna pattern calculation at the location of user i, (x_i, y_i, z_i) from the location of ap_j, (x_j, y_j, z_j)

- **Interference level ($Intf_{ij}$)**

The interference level for user i when it associates to ap_j ($Intf_{ij}$) is the summation of interfering signals from other APs in the vicinity.

$$Intf_{ij}\,(dBm) = 10\,log_{10}\left(\sum_{\forall k \in A/\, j \neq k}\left(\eta_{jk}.log_{10}^{-1}\left(\frac{P_{R_{ik}}}{10}\right)\right)\right) \tag{3.2}$$

η_{jk} is the percentage of interfering power that the signal operating at the channel f_k dissipates to the signal operating at the channel f_j due to the overlapping of the frequency spectrum of the channels. Consider the frequency spectrum of the 802.11b and 802.11g as shown in Figure 1.2. Channels f_j and f_k are non-overlapping channels and do not interfere with each other (i.e., $\eta_{jk} = 0$) if the channel separation (channel distance) is at least five (i.e., $|f_j - f_k| \geq 5$). Channels f_j and f_k are overlapping channels and interfere with each other if the channel distance is less than five (i.e., $|f_j - f_k| < 5$). In such a case, $0 < \eta_{jk} \leq 1$. In the case that APs are assigned the same frequency channel (co-channel) and the channel distance is zero ($f_j = f_k$), they interfere with each other 100 % ($\eta_{jk} = 1$). In the case of partially overlapping-channels (i.e., $0 < |f_j - f_k| < 5$), η_{jk} is computed from the ratio of the overlapping spectrum as in the approach used in the OPNET simulation [65]. For example, consider f_j = channel 1 (the frequency range = 2.401 – 2.423 GHz) and f_k = channel 2 (the frequency range = 2.406 – 2.428 GHz). The bandwidth of each channel is 22 MHz. In this case, the channel separation = 1 and the overlapping bandwidth is 2.423

GHz – 2.406 GHz = 17 MHz. Thus, the ratio of the overlapping spectrum, η_{jk}, is $\frac{2423-2406\ MHz}{22MHz}$ = 0.77. When the channel separation is 2, 3, and 4, the overlapping bandwidth is 12, 7, and 2 MHz, respectively and η_{jk} = 0.54, 0.31, and 0.1, respectively.

Note that in equation (3.2), the received signal ($P_{R_{ik}}$) is in units of dBm. In order to sum the signals, dBm has to be converted to miliWatts by dividing $P_{R_{ik}}$ by 10 and taking anti $\log_{10}$. After summing, the obtained signal is converted back to dBm by taking $\log_{10}$ and multiplying by 10.

- **Average obtainable user data rate (r_i^t)**

Since network usage characteristics and traffic correlate with locations where users situate, we incorporate such characteristics in the demand-based WLAN design through two parameters: the user activity level (α_t) and the average data traffic (R_t) generated by active users. When employing the 802.11 medium access protocol (CSMA/CA), the AP's capacity and the available user data rate reduce with the increasing number of active users competing for the AP's data channel [10, 66, 67]. Satisfying the average user data rate requirement by deploying extra APs in the service area may not yield higher network capacity if the extra APs increase interference. In the proposed WLAN design model, an adequate number of APs are deployed to accommodate the expected amount of traffic and provide sufficient average data rate to the target wireless users while limiting the interference level in the service area. Given the expected number of active wireless users associating to an AP, the available user data rate can be estimated by using an 802.11 capacity analytical model [10, 66, 67]. In the calculation of the average available user data rate and the capacity analytical model, we use the following notation:

Notation:

U_j^t = Set of all prospective wireless users located in sub-area type t that are within the radio signal coverage of ap_j and associate to ap_j

$$\left|U_j^t\right| = \sum_{i \in U} \sum_{j \in A} d_{ij}^t$$

m_j^t = Set of active wireless users located in sub-area type t that are active in data transfer activities and sharing capacity of ap_j. The average number of active users in the set m_j^t is the fraction of users in the set U_j^t that are expected to be active and can be estimated by applying the parameter α_t (user activity level) (i.e., $\left|m_j^t\right| = \alpha_j \left|U_j^t\right|$)

m_j = Set of all active wireless users that are sharing capacity of ap_j

= Union of all active users from all sub-areas that can communicate to ap_j

= $\bigcup_{\forall t \in T} m_j^t$

Λ = Set of all active wireless users in the network = $\bigcup_{\forall j \in A} m_j$

pkt_t = Average packet length (bits) of data traffic from the user in sub-area type t

τ_i^t = Average overall transmission time (μsec) that it takes to successfully transmit a packet from user i located in sub-area type t

$\tau_{overhead}$ = Transmission time according to the CSMA/CA protocol overhead (μsec)

$\tau_{compete}$	=	Estimated time spent in the contention period when a group of m_j active users are contending for the data channel at ap_j (μsec)	
τ_{pkt}	=	Packet transmission time (μsec)	
$P_c(m_j)$	=	proportion of collisions experienced by each packet successfully acknowledged at the MAC layer when a group of m_j active users are contending for the data channel at ap_j ($0 \le P_c(m_j) < 1$)	
DIFS	=	Distributed Interframe Space	= 50 μsec
$PLCP_{preamble}$	=	Preamble part of the Physical Layer Convergence Protocol	= 72 μsec
$PLCP_{header}$	=	Header part of the Physical Layer Convergence Protocol	= 24 μsec
SIFS	=	Short Interframe Space	= 10 μsec
ACK	=	Acknowledgement	= 10 μsec
SLOT	=	Time slot	= 20 μsec
$CW_{\min}$	=	Contention window	= 240 bits
CRC	=	Cyclic Redundancy Check	= 32 bits
MAC_{header}	=	Average medium access layer header	= 240 bits

Most CSMA/CA capacity analytical models assume asymptotic conditions where terminals always have a packet ready for transmission [10, 66, 67]. The asymptotic condition of the capacity analytical model is applied only to active terminals that have a packet ready for transmission. Therefore, $|m_j|$ terminals compete for the radio channel. The long term channel access probability of CSMA/CA is equal for all terminals and an average available user data rate is computed by assuming that the wireless terminals alternate transmissions with the probability of collisions during contention for channel access using the model of Heusse et al. [67]. We adopt the capacity analytical model presented by Heusse et al. [67] to compute the average available user data rate (r_i^t), which is the ratio between the average packet size of user *i* and the time it takes for successful transmission, as written in equation 3.3. Figure 3.3 shows the CSMA/CA MAC variables, the contention period and the successful transmission period. Note that the successful transmission period accounts for a constant overhead time ($t_{overhead}$, as written in equation 3.5), time in contention procedures ($t_{compete}$, as written in equation 3.6), and a package transmission time (t_{pkt}, as written in equation 3.7).

$$r_i^t \text{ (bps)} = \frac{pkt_t}{\left(\sum_{\forall i \in m_j} \tau_i^t\right)\left(1 + P_c(m_j)\right)} \qquad \text{for } \forall j \in A \qquad (3.3)$$

Where

$$\tau_i^t \quad (\mu\text{sec}) \quad = \tau_{overhead} + \tau_{compete} + \tau_{pkt} \qquad (3.4)$$

$$\tau_{overhead} \quad (\mu\text{sec}) \quad = DIFS + 2 \times PLCP_{preamble} + 2 \times PLCP_{header} + SIFS + ACK \qquad (3.5)$$

$$\tau_{compete} \quad (\mu\text{sec}) \quad = SLOT \times \frac{1 + P_c(m_j)}{2|m_j|} \times \frac{CW_{\min}}{2} \qquad (3.6)$$

$$\tau_{pkt} \quad (\mu\text{sec}) \quad = \frac{pkt_t + MAC_{header} + CRC}{C_j} \qquad (3.7)$$

$$P_c(m_j) = 1 - \left(1 - \frac{1}{CW_{\min}}\right)^{|m_j| - 1} \qquad (3.8)$$

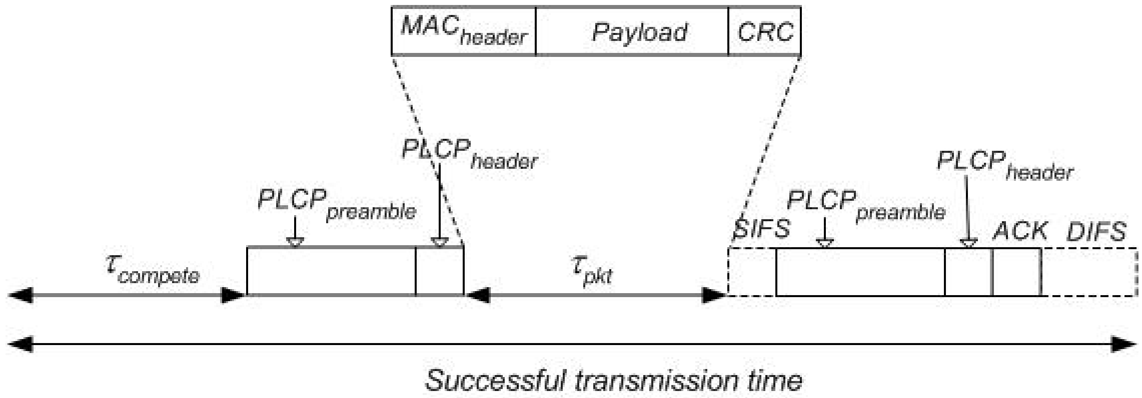


Figure 3.3 Structure of a successful transmission time

The calculation of the received signal strength ($P_{R_{ij}}$), the interference level ($Intf_{ij}$), and the average user data rate (r_i^t) described above are incorporated into the set of constraints in the proposed WLAN design model.

3.4.5 Constraint Satisfaction Problem Formulation

This section presents a mathematical formulation of the demand-based WLAN design problem. The problem is formulated as a Constraint Satisfaction Problem (CSP), with prescribed requirements for a finite number of variables with a given set of possible values (called domains) that can be assigned to the variables [51]. A constraint specifies which value tuples are allowed for a certain subset of all the variables. The solution of a CSP is any instantiation of all the variables that satisfy all constraints simultaneously.

The CSP for the demand-based WLAN design model imposes restrictions and network requirements into the design process. It can be defined by the triple (V, D, C), where

V = $\{n, p_j, f_j, d_{ij}^t, g_{hj}, (x_j, y_j, z_j)\}$ denotes a set of variables of the design problem

D = $\{D_n, D_p, D_f, D_d, D_g, D_{(x,y,z)}\}$ denotes a set of finite domains associated with each variable

C = {C1, C2, C3, C4, C5, C6, C7, C8, C9} denotes a set of constraints

The number of APs used in the network, *n*, is determined by the construction phase of the solution technique (as described in section 4.2). For the specified number of APs (*n*), the APs' parameters (e.g., the location, frequency channel, and power level) are determined such that all constraints in set **C** are satisfied. If all constraints cannot be satisfied for the initial number of APs, *n*, the Add-AP phase (described in section 4.6) is applied to place additional APs in the network.

The constraints in the CSP for the demand-based WLAN design model include:

C1: $$\sum_{\forall j \in A} d_{ij}^t = 1 \quad , \forall i \in U \tag{3.9}$$

C2: $$d_{ij}^t \left(P_{R_{ij}} - P_{Rthreshold}\right) \geq 0 \quad , \forall i \in U, \forall j \in A \tag{3.10}$$

C3: $$d_{ij}^t \left(P_{R_{ij}} - Intf_{ij} - SIR_{threshold}\right) \geq 0 \quad , \forall i \in U, \forall j \in A \tag{3.11}$$

C4: $$r_i^t > R_t \quad , \forall i \in U \tag{3.12}$$

C5: $$\sum_{\forall j \in A} g_{hj} \geq 1 \quad , \forall h \in G \tag{3.13}$$

C6: $$g_{hj} \left(P_{R_{hj}} - P_{Rthreshold}\right) \geq 0 \quad , \forall h \in G, \forall j \in A \tag{3.14}$$

C7: $$g_{hj} \left(P_{R_{hj}} - Intf_{hj} - SIR_{threshold}\right) \geq 0 \quad , \forall h \in G, \forall j \in A \tag{3.15}$$

C8: $$d_{ij}^t \in \{0,1\} \quad , \forall i \in U, \forall j \in A$$

C9: $g_{hj} \in \{0,1\}$ $, \forall h \in G, \forall j \in A$

The defined set of constraints imposes the design requirements including radio signal quality and data rate capacity criterion into the network planning process. Constraints C1-C3 and C5-C7 address the radio signal quality requirement while constraint C4 addresses the data rate capacity requirement.

The set of constraints C1-C3 ensure that the prospective wireless users in the service area can connect to the WLAN. That is, satisfying C1-C3 simultaneously results in each demand node having adequate received signal strength and SIR level such that wireless data transfer can take place. Constraint C1 requires that each wireless user associates to one AP and only one AP. The decision variable d^t_{ij} is equal to one if the received signal strength that user i received from the ap_j (P_{Rij} in dBm) and the SIR level with respect to the ap_j (the received signal strength (P_{Rij} in dBm) less the interference level ($Intf_{ij}$ in dBm)) meet the receiver sensitivity threshold ($P_{Rthreshold}$) and the SIR threshold ($SIR_{threshold}$) as specified by C2 and C3, respectively; d^t_{ij} is equal to zero otherwise.

Constraint C4 ensures that the average data rate available to wireless user i which is a type t user (r_i^t) is greater than the specified user data rate (R_t). The 802.11 capacity analytical model and the user activity pattern correlated with the type of sub-areas where users locate are incorporated in this constraint to estimate the average data rate that the active wireless user can obtain [10, 67].

The set of constraints C5 – C7 ensure that the radio signal is available throughout the predefined usable area of the service region. To assess the signal quality in the service area, the received signal strength and the SIR level are tested at all signal test points (STPs) specified by the grid spacing granularity. Constraint C5 specifies that each STP must be able to receive signals from at least one AP. It allows overlapping of the APs' coverage areas. The decision variable g_{hj} is equal to one if the received signal strength at the STP h transmitted from the ap_j (P_{Rhj} in dBm) and the SIR level with respect to the ap_j (i.e., P_{Rhj} - $Intf_{hj}$) meet the received sensitivity threshold ($P_{Rthreshold}$) and the SIR threshold ($SIR_{threshold}$) as specified by C6 and C7, respectively; g_{hj} is equal to zero otherwise.

Constraints C8 and C9 specify that variable d_{ij}^{t} and g_{hj} are binary {0, 1} variables, respectively.

3.5 Solution Technique for the Demand-based WLAN Design

The demand-based WLAN design approach aims to find any single feasible network configuration that satisfies the network design criteria. Consider the complexity of the CSP for a problem with n number of variables in which each variable consists of a candidate values. There are altogether a^n possible combinations of n-tuples (candidate solutions). In the worst case, the search may have to explore the solution space exhaustively to identify a feasible solution if any exists. The complexity of the CSP module is $O(a^n)$, which increases exponentially with the number of variables n. Prommak et al. [68] show that the computational time of an exhaustive search of the CSP WLAN design solution space increases exponentially with the number of variables. The next chapter presents an efficient heuristic solution technique that is developed in this dissertation to solve demand-based WLAN design problems.

3.6 Summary

This chapter describes the demand-based WLAN design approach in detail. The developed demand-based WLAN design approach incorporates two essential design criteria: radio signal coverage and the data rate capacity. The demand node concept and the busy-hour user population are applied to represent user distribution and traffic demand in the service area. Typical WLAN network usage characteristics and user activity level are incorporated into the network design model which is formulated as a constraint satisfaction problem (CSP). Due to the complexity of the resulting CSP network design model, a heuristic solution technique is proposed to efficiently solve the demand-based WLAN design problem in reasonable time. The next chapter describes the heuristic solution technique in detail.

4.0 A HEURISTIC SOLUTION TECHNIQUE FOR DEMAND-BASED WLAN DESIGN

Demand-based WLAN design aims to devise a network configuration that can provide signal coverage and accommodate traffic demand in a target service area. Network requirements are incorporated into the design model through application of the constraint satisfaction problem (CSP) formulation presented in chapter three. This chapter presents a heuristic solution technique that we developed to solve the demand-based WLAN design problem. The heuristic solution technique will determine the number of APs required in the service area and their parameters, including location, power level and frequency channel in the WLAN configuration. Section one presents the overall framework of the heuristic solution technique, consisting of five phases: the construction phase, the frequency channel assignment (FCA) phase, the constraint violation reduction (CVR) phase, the intensification phase and the add-AP phase. The following sections of this chapter describe the detailed implementation of each phase.

4.1 A Framework for Solving the WLAN CSP

Demand-based WLAN design can be formulated as a constraint satisfaction problem (CSP) as presented in chapter three. The objective of the design is to identify a feasible solution that can satisfy the specified network requirements. Since the complexity of the CSP design model increases exponentially as the problem size (or number of variables) increases, we developed a heuristic solution technique that can efficiently explore the search space and find a feasible solution within a reasonable amount of computational time.

The overall framework of the heuristic solution technique is shown in Figure 4.1. There are two input components for the main solution technique. The first input involves the physical description of the service area (e.g., building size, location and composition of partitions, etc.) and the user traffic information (e.g., user location density, user activity level and required average data rates). The second input specifies models (functions) and parameters for fundamental calculations necessary to the solution technique, including the path loss model, the 3-D antenna radiation model, and the capacity analytical model. The parameters include the necessary information to formulate and compute the constraints in the CSP (e.g., SIR threshold, received signal strength threshold, path loss exponent, etc.). The output from the solution technique is a WLAN configuration that meets the network design requirement. It specifies the number of APs required and their parameters, including location, power level, and frequency assignment.

The main body of the solution technique consists of five phases; the construction phase, the frequency channel assignment (FCA) phase, the constraint violation reduction (CVR) phase, the intensification phase and the add-AP phase. The interconnection of the phases is shown in Figure 4.1.

An overview of the solution technique process is as follows. The first two phases (the construction and FCA phases) aim to generate a good starting configuration that provides an estimated number of APs and their initial parameters. The construction phase (described in detail in section 4.2) involves two heuristics: the Area Covering Heuristic

(ACH) and the Demand Clustering Heuristic (DCH). The ACH involves estimating the number of APs that are required to provide radio signal coverage to the service area, while DCH deals with placing additional AP(s) in those parts of the service area where high traffic volume exists. Together, these two heuristics determine the initial locations and power levels of the APs. The FCA phase (described in detail in section 4.3) utilizes a simulated annealing method to determine the frequency channel assignments of the APs based on their initial locations and power levels as defined in the construction phase.

The CVR phase (described in detail in section 4.4) evaluates the initial network configuration by using evaluation functions described in section 4.4.3. If any design requirement constraint is violated, the CVR phase reduces the constraint violations by adjusting the locations and power levels of the APs by using tabu search and reassigning frequency channels by using the FCA simulated annealing method. If the CVR phase fails to produce a feasible network configuration satisfying all design constraints, the intensification phase (section 4.5) revisits good candidate solutions recorded during the CVR phase and performs the repairing process for each revisited solution. After the intensification phase, if a feasible network configuration is still not found, the add-AP phase (section 4.6) attempts to solve the problem by installing additional AP(s) in the service area.

The implementation of each phase of the solution technique is described in detail in the following sections.

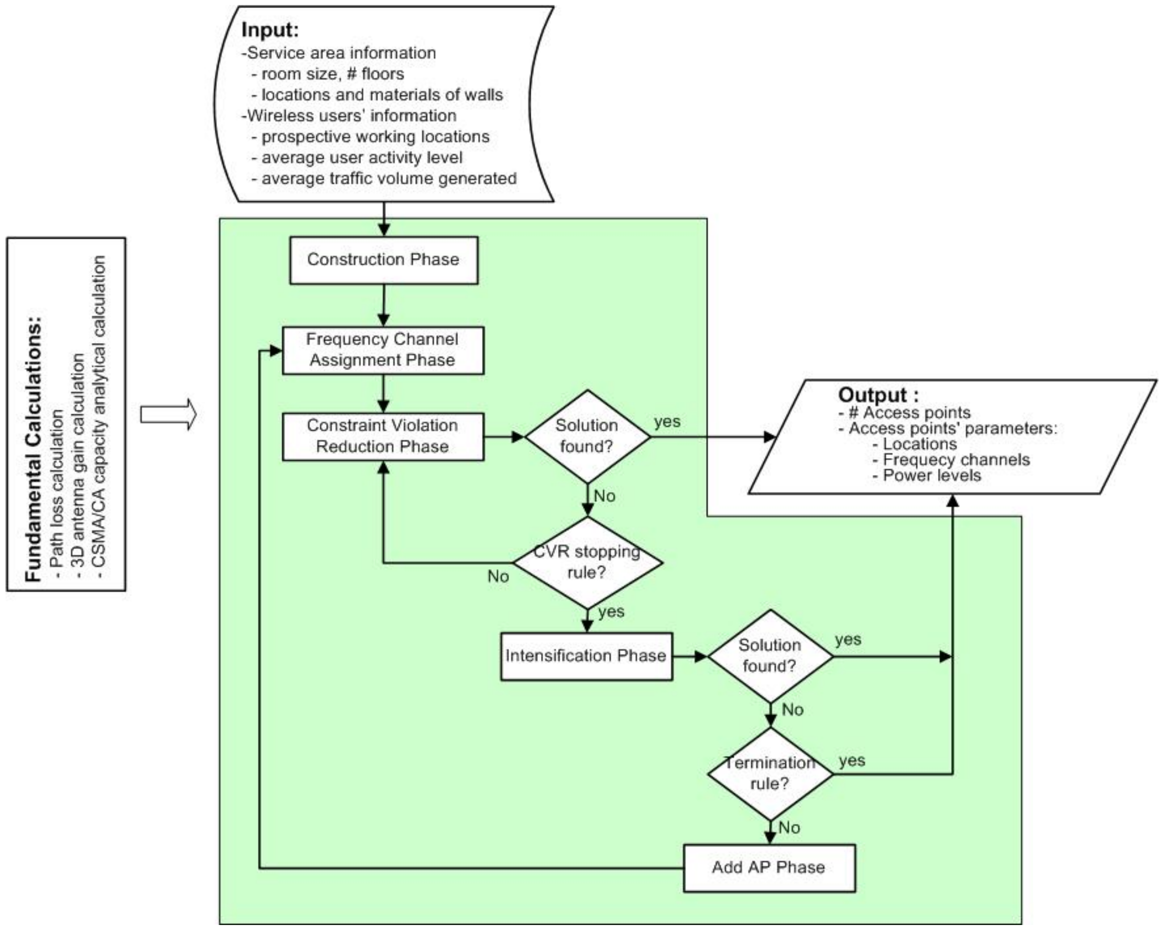


Figure 4.1 Framework of the solution technique

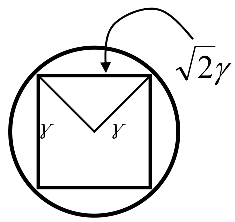
4.2 Construction Phase

For a given design scenario, the construction phase generates an initial solution which provides the number of APs as well as their locations and power levels. The construction phase is a combination of the Area Covering Heuristic (ACH) and the Demand Clustering Heuristic (DCH). The ACH is a modified version of the Initial Covering Heuristic (ICH) [24], which estimates the number of transmitters required in coverage optimized wireless network design. The DCH was developed based hierarchical clustering methods [69]. It clusters and represents a group of wireless terminals with a single point where a service facility should be placed. The ACH and DCH are explained in detail as follows.

4.2.1 Area Covering Heuristic (ACH)

The ACH aims to determine the initial number of APs for the given service area. Unlike the ICH in which the APs are uniformly distributed across the service region [24], the ACH determines location of APs based not only on the size of the service area but also on the distribution of traffic demand in the area.

Assessing the initial number of APs required for a given service scenario requires estimating the coverage area of each AP. Here we assume an isotropic antenna and approximate AP coverage area as a circle with a radius r computed using a path loss model for the given environment and the minimum required received signal strength. In order to allow an overlap between coverage of adjacent APs, we represent the access point coverage area as a coverage square (CS), which is the biggest square that can fit into the coverage circle as shown in Figure 4.2.



CS = $\sqrt{2}\gamma \times \sqrt{2}\gamma$

Figure 4.2 Coverage Square (CS) estimating an AP's coverage area

Given the operating frequency (f), the AP transmitting power (P_T), and the receiver sensitivity (P_R), γ is calculated from the equation (4.1).

$$\begin{aligned} P_T - L\big(f,(x_i,y_i),(x_j,y_j)\big) &= P_R \\ L(f,(x_i,y_i),(x_j,y_j)) &= L(d_0) + 10 n_0 \log\left(\frac{\gamma}{d_0}\right) + X_\delta \end{aligned} \tag{4.1}$$

where $L(d_0)$ is the path loss at reference distance d_0 , n_0 is the path loss exponent, and X_δ is a log-normally distributed random variable representing the shadow fading with standard deviation δ dB.

For example, consider an AP that operates at 2.462 GHz. A minimum received signal strength requirement is P_R = -80 dBm [70], given the highest power level P_T = 24 dBm, d_0 = 1 meter [12], n_0 = 3.3 [71], δ = 3.5 dB [72], and the corresponding X_δ = 5.75 dB for 95% coverage availability at the edge of the cell [62]; solving equation (4.1); we obtain γ = 23 m. Thus, the coverage square is CS = $\sqrt{2}(23)\times\sqrt{2}(23)$=32.5×32.5 m^2.

The resulting CS is used in the ACH, a bottom up process which starts from the first floor and progresses through the top floor. Let $\overline{U}$ be a set of wireless users who are not associated to any AP and U represents the set of all wireless users in the service area. Initially, $\overline{U} = U$. A sufficient number of CSs are placed on each floor to cover the entire area. Typically, one AP is placed in each CS. However, there is no need to install any AP in those CSs where low traffic demand exists and signal coverage is available from another floor. In each CS, the location of the AP and its power level are determined as follows. If low traffic demand exists, the AP is located in the middle of the CS and assigned the highest power level. All wireless users in the CS are assigned to the AP and removed from $\overline{U}$. In a CS where traffic demand is high (i.e., the required user data rate of all wireless users in the CS cannot be satisfied by one AP), wireless users are selected and associated to an AP located at the center of gravity (CoG) of the cluster of the selected wireless users. The cluster of users is formed such that each user in the cluster can obtain the required average data rate.

The cluster of wireless users and the corresponding CoG are formed as follows:

- To form a cluster of wireless users in a CS where high traffic demand exists:

i. Select the first wireless user closest to the left corner of the CS (denote it as u_1). Sort the rest of the wireless users in the CS in an increasing order according to their distance from u_1. Call the resulting list *sorted_List*.

ii. Add the first wireless node of the *sorted_List* to the cluster and compute the CoG of the cluster.

iii. If the required user data rate can be satisfied, remove the wireless user who was just been added to the cluster from the set $\overline{U}$. Otherwise, go to step v.

iv. If there remain unassigned wireless users in the CS, go to step ii.

v. The unassigned wireless terminals in the CS will be taken care of in the DCH process. For the users in CS that are assigned an AP, the AP is placed at the resulting CoG derived by equation 4.2 shown below. The power level to be assigned to the AP is determined from the Euclidian distance measured from the location of u_1 to the resulting CoG. This distance is called *distance_to_CoG* and is determined in the following ways:

If *distance_to_CoG* < Υ_0, power level = P_0

If *distance_to_CoG* < Υ_1, power level = P_1

…

If *distance_to_CoG* < Υ_{max}, power level = P_{max}

where Υ_0, Υ_1, … Υ_{max} are computed by using equation 4.1 as illustrated previously.

Υ_0 is the resulting radius of the coverage area when $P_T = P_0$

Υ_1 is the resulting radius of the coverage area when $P_T = P_1$

…

Υ_{max} is the resulting radius of the coverage area when $P_T = P_{max}$

Note that at the end of the ACH process there may be unassigned wireless users remaining in set $\overline{U}$. The DCH process will consider adding more APs to accommodate the traffic demand of those users.

- To derive the center of gravity (CoG):

Let $\left(x_{CoG}, y_{CoG}\right)$ denote the coordinates of a particular CoG.

$$x_{CoG} = \frac{\sum_{i \in Cluster(j)} x_i}{n_{(j)}} \text{ and } y_{CoG} = \frac{\sum_{i \in Cluster(j)} y_i}{n_{(j)}} \tag{4.2}$$

where (x_i, y_i) is the coordinates of wireless user i who is in cluster j, and $n_{(j)}$ is the number of wireless users in cluster j.

4.2.2 Demand Clustering Heuristic (DCH)

In an area with a high user density and traffic volume where some wireless nodes are unassigned by the ACH process (i.e., the wireless terminals in the set $\overline{U}$), the DCH attempts to cluster those terminals and derive a center of gravity for each cluster where an additional AP will be placed. Unlike the clustering procedures used in the ACH, in which the cluster of wireless nodes is formed within the CS, here the cluster of wireless nodes can be formed across multiple CS boundaries. The DCH is a bottom-up process, progressing from the first through the top floors.

For each floor i:

i. Determine the set of unassigned wireless nodes on floor i, denoted by $\overline{U_i}$.

ii. Try to associate wireless nodes in $\overline{U_i}$ to existing APs with residual capacity in the adjacent floor. Remove the assigned wireless nodes from $\overline{U_i}$.

iii. Select an unassigned wireless node in $\overline{U_i}$, which is closest to the left corner of floor i (denoted as u_1). Sort the remaining unassigned wireless nodes in $\overline{U_i}$ in increasing order according to their distance from u_1. Store this as a *sorted_List*.

iv. Add the first wireless node of the *sorted_List* to the cluster and compute the CoG of the cluster. Let *distance_to_CoG* be the distance from the u_1 to the CoG. Let γ_{max} be the radius of the coverage area of an AP placed at CoG when operating at the highest power level. If *distance_to_CoG* < γ_{max}, add the next wireless node in the *sorted_List* to the cluster and recompute the CoG. The wireless nodes in the clusters are removed from $\overline{U_i}$.

v. The clustering process stops when the *distance_to_CoG* ≥ γ_{max} or the obtainable data rate of the wireless nodes in the cluster is lower than the required data rate.

vi. The resulting CoG is the location to place an additional AP. The power level to be assigned to the AP is determined from the *distance_to_CoG* as described previously.

vii. If the $\overline{U_i}$ is not empty, return to step iii.

viii. If the $\overline{U_i}$ is empty, continue onto the next floor.

4.2.3 Example of the Construction Phase

Figure 4.3 illustrates an example of the construction phase. A single floor service area and its wireless node distribution is shown in Figure 4.3a. The first CS is placed in the lower left corner of the service area, as shown in Figure 4.3a. Figure 4.3b shows the service area covered with the minimum number of CSs (in this case nine). Next, the APs are placed and their power levels are defined. This step addresses the traffic distribution in the service area and incorporates the AP capacity analytical model [10, 67] to estimate both the average data rate available to users and the effective capacity of the AP.

Associate all wireless nodes located in each CS to the AP located in the middle. If all wireless nodes can obtain an average data rate as specified in the traffic profile, the AP is assigned with the highest power level as shown in Figure 4.3c. If the required average data rate cannot be achieved due to a high density of users or high traffic demand (see the second CS shown in Figure 4.3d), the demand clustering technique from section 4.2.2 is applied to derive the power level and location for APs needed in that CS. In Figure 4.3d, the middle power level is assigned to the AP in the second CS, which is located at the CoG of the cluster of selected wireless nodes. After proceeding through all CSs, Figure 4.3e presents the result from the ACH process. We can see that in areas of high traffic density, some wireless nodes are left unserved. The DCH heuristic is then applied to place additional APs in those areas. Figure 4.3f shows the results of the construction phase providing the initial number of APs, their locations and their power level assignment for the given network design scenario.

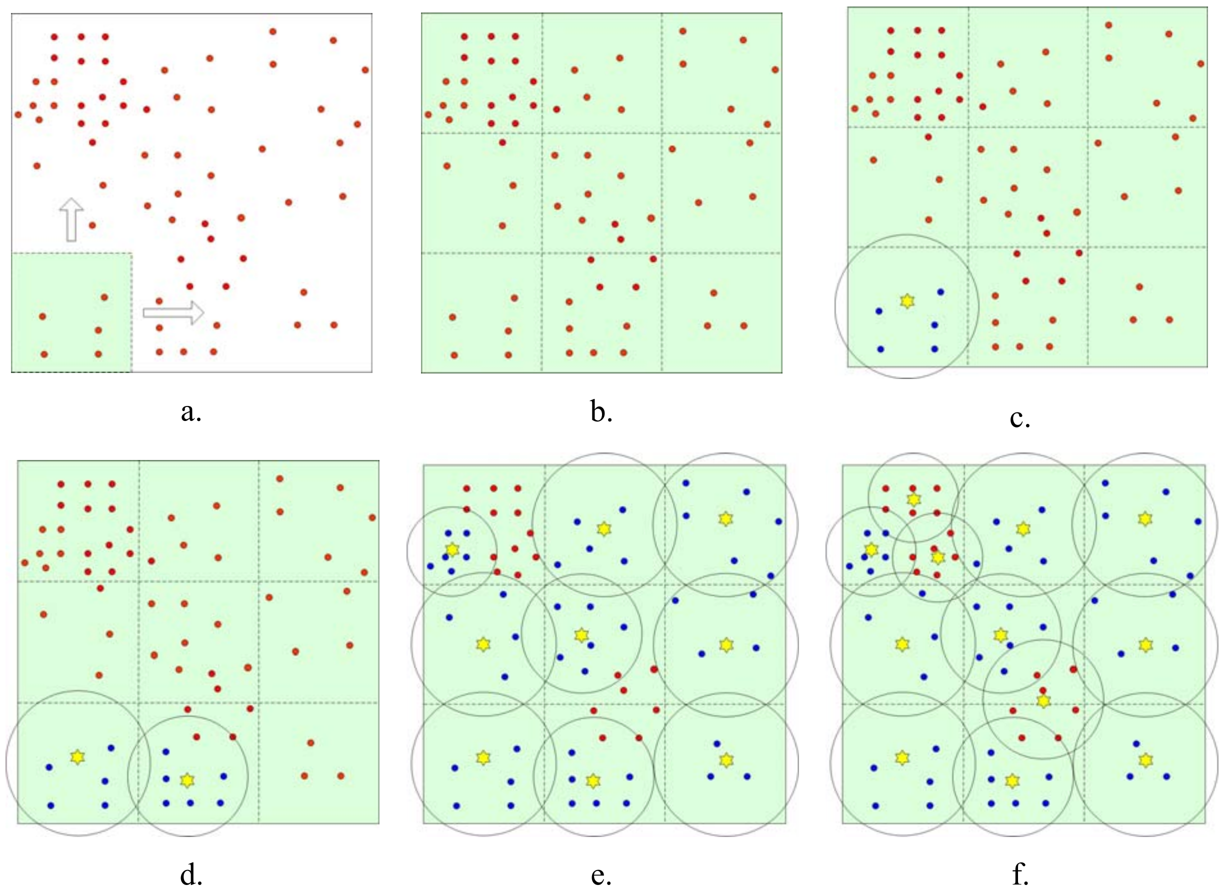


Figure 4.3 Construction phase procedures

4.3 Frequency Channel Assignment Phase

The frequency spectrum available for WLAN operations is limited. In North America, of a total of eleven frequency channels in the 2.4 GHz band, just three non-overlapping channels are allocated for 802.11b and 802.11g operation [8], while twelve non-overlapping channels in the 5 GHz band are allocated for 802.11a operation [7]. Due to the limited number of WLAN operating channels available, frequency channels need to be carefully assigned to APs so that the resulting network can maintain an adequate signal to interference ratio (SIR).

This section describes the frequency channel assignment (FCA) procedure in the developed solution technique for demand-based WLAN design. Inputs to the FCA phase are the APs' locations and their power levels as obtained during the construction phase (section 4.2). The output from this phase is the frequency channel assignment for each AP in the network.

The FCA problem in demand-based WLAN design is a fixed-spectrum FCA problem [73]. It involves assigning frequency channels from a limited set of frequency channels to a given set of transmitters (i.e., APs) such that the signal strength of the desired signal is greater than the interference signal evaluated at all signal test points (STPs) in the coverage area.

In general, the FCA problem is described as a graph coloring problem [74] in which each vertex of the graph represents a transmitter in the network and each edge of the graph represents a constraint on the frequency channel separation between two APs at the end of each edge. The FCA phase is divided into two steps: generating channel separation constraints and assigning frequency channels to APs based on the channel separation constraints. Each step is described in detail as follows:

4.3.1 Generating Channel Separation Constraints

Let graph $\psi = (V, E)$ represent a WLAN in which $V = \{v_1, v_2, \ldots, v_n\}$ is the set of vertices representing APs in the set $A = \{ap_1, ap_2, \ldots, ap_n\}$ and E = $\{(v_i,v_j)/\ v_i,v_j \in V)$ is the set of edges representing channel separation constraints between all pairs of APs.

In FCA, the channel separation constraints are used to specify channel distances between all pairs of APs in the network. The constraints are represented by a matrix called the channel separation matrix, denoted by B. Let $B = (b_{ij})^{n' \times n'}$ where n' is the number of vertices of the graph, element b_{ij} with $i \neq j$ representing the minimum channel separation required between vertices v_i and v_j such that the network can avoid interference caused by other APs in the vicinity. This section explains the process of generating the channel separation matrix B.

We adopt the procedure to generate the channel separation matrix B explained by L.E. Hodge et al. [75]. The procedure begins by determining which AP in the network yields the highest received signal strength at each signal test point (STP) $g \in G$. Let *tuned_AP[g]* denote the AP that yields the highest received signal strength to the STP g. For each STP $g \in G$, the signals from all other APs are considered in turn as interfering signals and a frequency separation constraint is generated (or strengthened) where necessary. Figure 4.4 shows the pseudo code of the process to generate the channel separation matrix B, which is used to store the *channel_distance* (b_{ij}) required between

ap_i and ap_j for $\forall\ i\ ,j\ \in A$ and $i \neq j$. Initially, all elements in the matrix contain zeros. The matrix B is derived such that for each STP g $\in G$, the SIR level (the signal strength from the *tuned_AP[g]* less the signal strength from the interfering AP) meets the required threshold. Note that the $distance_{max}$ (in Figure 4.4) denotes the channel separation between each pair of consecutive non-overlapping channels. For 802.11b and 802.11g, the non-overlapping channels include channels 1, 6, and 11. Thus, the $distance_{max}$ = 5. For 802.11a, the value of $distance_{max}$ is set to 1 because in 802.11a, there are no overlapping channels. Thus, the $distance_{max}$ between two consecutive channels is one.

```
Initialize frequency channel of each AP (let all APs operate at the same channel)
Initialize b_ij (written as b[i][j]) such that each element contains 0, for all AP i and j ∈ A

For each signal test point g ∈ G
  Determine tuned_AP[g] = AP that yields the highest signal strength at g
End For

/* Generate the channel separation matrix, B */
For each signal test point g ∈ G
  For each AP j ∈ A where j≠ tuned_AP[g]
      P_tuned = signal strength received at g from the tuned_AP[g] (in dBm)
      P_interfering = signal strength received at g from AP j (in dBm)
      SIR = P_tuned − P_interfering  (in dB)
      channel_distance = 0
      newP_interfering = P_interfering (in dBm)

      While(SIR < SIR_threshold & channel_distance < distance_max)
          channel_distance = channel_distance +1

          newP_interfering = 10log_10(percent_interfering[channel_distance]* 10^(P_interfering/10) )  (in dBm)
          SIR = P_tuned − newP_interfering (in dB)
       End While

      If(channel_distance > b[tuned_AP[g]][j])
         b[tuned_AP[g]][j] = channel_distance
  End For
End For
```

Figure 4.4 Pseudo code for generation of the channel separation matrix

As an example, consider a WLAN design for a two-floor service environment. Assume that the construction phase produces a WLAN setting as shown in Figure 4.5a where two and three APs are located on the first floor and the second floor, respectively. ap_1, ap_2, and ap_4 are assigned a transmit power of 24 dBm while ap_3 and ap_5 are assigned transmit power of 15 dBm. Figure 4.5b illustrates the resulting channel separation matrix for the WLAN setting of Figure 4.5a. In the matrix B in Figure 4.5b, the element $b[1][2]$ = 3 means that the minimum channel separation required between ap_1 and ap_2 is three channels. Figure 4.5c shows the graph based on the matrix B. Note that for elements of B that equal zero, no edge exists between the corresponding vertices in the graph. This indicates that the corresponding pair of APs can use the same frequency channel.

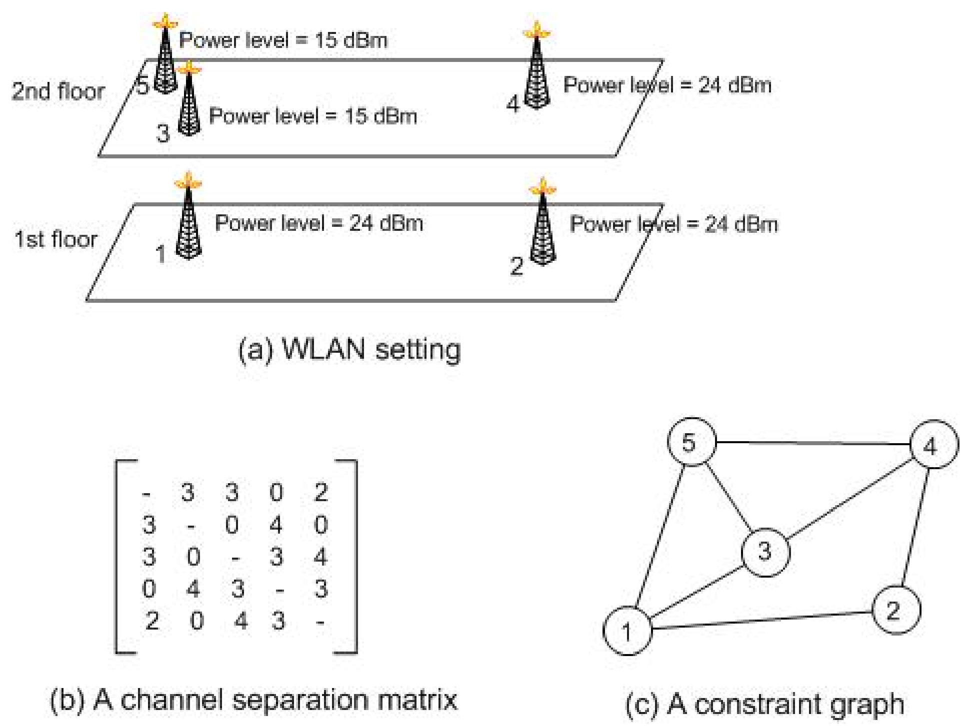


Figure 4.5 Deriving a channel separation matrix and a network graph

4.3.2 Assigning Frequency Channels to APs

The channel separation constraints explained in the previous section are used to restrict the frequency channels assigned to APs in the network. The FCA problem in WLAN design can be described as follows. Given a set of frequency channels available in the frequency spectrum allocated for WLAN operation, the frequency channels are assigned to APs such that the channel separation constraints specified by the matrix *B* are satisfied. If overall satisfaction is impossible, the goal is to minimize any violations of the channel separation constraints.

As defined in chapter 3, the notation reviewed below is used in the mathematical formulation of the FCA problem.

Notation:

f_j = Frequency channel assigned to AP *j*

D_f = A set of consecutive integers specifying available frequency channels for the WLAN design = $\{F_1, F_2, \ldots, F_k\}$

FCA problem formulation:

Minimize $$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} max\left(0, b_{ij} - | f_i - f_j |\right) \tag{4.3}$$

Subject to: $$f_i, f_j \in D_f, \text{ for } \forall\ i, j \in A \tag{4.4}$$

Several techniques have been developed to solve the FCA problem. Aardal et al. [76] provide a survey on existing solution approaches which include; graph theory [77], integer linear programming [78], greedy techniques [79], and meta-heuristic approaches such as tabu search [79, 80], simulated annealing [79, 81], and genetic algorithm [79, 82].

Here, we adopt the simulated annealing (SA) method to solve the FCA problem for this WLAN design because of its simplicity and effectiveness as reported by Dunkin and Allen, and Thiel et al. [73, 74]. The following sections provide an overview of the SA algorithm and describe in detail the implementation of SA for the WLAN FCA problem.

4.3.2.1 Overview of Simulated Annealing

Simulated annealing (SA) is a variation of a hill-climber heuristic search approach. Unlike a greedy hill-climber heuristic which only moves to a solution that improves the objective function, SA allows the search to move to non-improving solutions (those that lead to a worse objective value) with a certain probability. This allows the SA search to avoid being trapped at local optima. The probability of accepting non-improving moves is controlled by a parameter which is analogous to the temperature in the annealing schedule at which liquids freeze into a minimum energy crystalline structure [83].

In order to make use of the simulated annealing method for solving mathematical optimization problems, the following elements must be provided:

- Solution representation: A description of a possible solution in the search space. This usually involves some configuration of parameters such as $\boldsymbol{X} = (x_1, x_2, \ldots, x_n)$ that represent a solution.

- Cost function: A function that maps each solution into a value representing its cost (analogous to the energy of the system in the thermodynamic case).
- Neighborhood: A set of solutions obtained by perturbing the current solution. One example of a perturbing technique is to randomly change any element x_i of the current solution while the rest remain unchanged.
- Control temperature *CT* and an annealing schedule: The annealing schedule simulating the physical process of a cooling system indicates how *CT* is lowered as the search progresses.
- Termination criterion: The algorithm stops when the termination criterion is satisfied. Generally the criterion is represented by a minimum value for the temperature or by a maximum number of consecutive iterations carried out without an improvement in the obtained solution.

4.3.2.2 Simulated Annealing Implementation for WLAN Frequency Channel Assignment

In order to apply the SA algorithm to solve the FCA problem for the demand-based WLAN design, the following components are defined:

- **Representation of the frequency channel assignment**

A frequency assignment is represented by using an array $\boldsymbol{F} = [f_1, f_2, \ldots, f_n]$ where n = the total number of APs in the network and f_j = the frequency channel assigned to AP j which is chosen from a set of available frequency channels given in the domain D_f.

- **Cost function**

The cost function, Θ, takes into account those factors involved in minimizing violations of the channel separation constraints. Such factors include the number of violated constraints ($n_{violated}$) and the sum of the amount by which each constraint is violated. The cost function, Θ, is written as follows:

$$\Theta = n_{violated} + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(max\left(0, b_{ij} - | f_i - f_j |\right)\right) \tag{4.5}$$

Note that even though both terms in the cost function are in different units, they are not normalized. The preliminary tests showed that normalized cost function did not yield good results in solving FCA problem; most of the time the search could not converge. The reason is that the normalized cost values were too small and caused premature freezing state of the annealing schedule, trapping at local optima. On the other hands, the non-normalized cost function in equation (4.5) yielded the search that could perform better than that using the normalized cost function; the preliminary tests showed that the search could converge to optimal solutions.

- **Initialization of frequency channel assignment**

An initial assignment is created by a greedy method that iteratively selects an AP and then assigns a feasible frequency to it. The selection follows a list of APs which are sorted in decreasing order according to the number of channel separations between the selected AP and all other APs. This is because we want to assign frequency channels to APs subject to many constraints before assigning to those with few constraints. The frequency channel to be assigned to the AP is selected so that the assignment avoids or minimizes channel separation violation of prior assignments.

- **Generation of new frequency channel assignment**

To explore the search space of the FCA problem, the SA algorithm generates a new frequency channel assignment ($\boldsymbol{F_{new}}$) based on the current $\boldsymbol{F}$, hoping that either $\boldsymbol{F_{new}}$ will be better than $\boldsymbol{F}$ or that it will lead to a better solution in future iterations of the search.

To produce $\boldsymbol{F_{new}}$, we define the neighbor solutions of the current $\boldsymbol{F}$ as those assignments where one element of the array $\boldsymbol{F}$ is altered. This can be done by randomly selecting one AP whose frequency channel will be changed and randomly selecting a new frequency channel for such APs. The move from the current assignment to the new assignment is accepted if it results in a reduction of the cost function. A non-improving assignment may be accepted with probability of $e^{-(\Theta(F_{new})-\Theta(F))/CT}$.

- **Annealing schedule**

The annealing schedule indicates how the parameter *CT* (analogous to temperature), which controls the probability of accepting non-improving moves, is lowered. In other words, it indicates how many random changes in the frequency assignment occur before *CT* is reduced as well as the extent to which it is reduced.

Here, we adopt a geometric annealing schedule as expressed in equation (4.6) in the SA algorithm for WLAN design because of its simplicity and effectiveness for graph coloring problems as presented by Costa et al. [81]. In the geometric annealing schedule, the temperature *CT* is reduced according to the parameter φ. The number of iterations at each temperature is specified by formula (4.7) where the initial value of Max_{iter} is set to n, the number of APs in the network. As the temperature *CT* decreases, the number of iterations (i.e., the number of random assignments tested at each temperature) increases. This scheme is adopted because as *CT* decreases and the process becomes more stable, a move tends to be accepted only when it yields improvement in the cost function. One can see that formula (4.7) allows a higher number of iterations (Max_{iter}) as the search progress. In this way, the search can explore the solution space thoroughly at a low value of *CT*. This annealing scheme was found to be effective in the convergence of the SA process as reported by Costa et al. [81].

$$CT = \varphi CT \tag{4.6}$$

$$Max_{iter} = \frac{Max_{iter}}{\varphi} \tag{4.7}$$

4.3.2.3 Procedures of the FCA SA Algorithm

Given a channel separation matrix obtained as described in section 4.3.1, Figure 4.6 illustrates the specific procedures of the SA algorithm when assigning frequency channels to the APs in a WLAN. The process begins by initializing parameters and generating an initial frequency channel assignment. Next, in each iteration, the algorithm generates a new frequency channel assignment ($\boldsymbol{F_{new}}$) from the current frequency channel assignment ($\boldsymbol{F}$). Let $\Delta cost$ be the different between the cost of $\boldsymbol{F_{new}}$ and that of $\boldsymbol{F}$. If the cost of $\boldsymbol{F_{new}}$ is less than or equal to the cost of $\boldsymbol{F}$ (i.e., $\Delta\text{cost} \leq 0$) then the new frequency channel assignment is accepted and becomes the current solution. Otherwise, the new assignment is accepted with a probability, P_{accept}, expressed as

$$P_{accept} = e^{-\frac{\Delta \cos t}{CT}} \tag{4.8}$$

The control temperature *CT* is varied as the search progresses. In the beginning, *CT* is high and, as the search proceeds, *CT* reduces. The manner in which the temperature *CT* reduces is specified by an annealing schedule (4.6). The search continues until the number of consecutive iterations without improvement reaches the *Max_noimprov* or until the temperature *CT* drops below CT_{min}.

```
Input static parameters: φ, Max_noimprov, CT_max, CT_min
Initialize CT = CT_max , Max_iter = n, iter = 0, count_noimprov = 0
Initialize frequency channel assignment F
Evaluate cost = Θ(F)

While CT > CT_min do
        While iter < Max_iter do
                Generate new assignment F_new from F
                Evaluate cost_new = Θ(F_new)
                Calculate Δcost = cost_new - cost
                If Δcost < 0 then
                        count_noimprov = 0
                        F = F_new
                        cost = cost_new
                Else
                        count_noimprov ++
                        If (count_noimprov > Max_noimprov) then End SA process
                        Else If random[0,1] < e^(-Δcost/CT) then
                                F = F_new
                                cost = cost_new
                        End If
                End If
                iter ++
        End While
iter = 0
CT = φCT
Max_iter = Max_iter / φ
End While
```

Figure 4.6 The simulated annealing procedure for frequency channel assignments

4.3.2.4 Static Parameters in FCA Phase

The following static parameters are determined from parameter tuning experiments and used in the SA process:

φ	=	the parameter that specifies how the temperature *CT* is reduced and how the number of iterations at each temperature step is increased as *CT* is reduced
Max_noimprov	=	the number of consecutive iterations that allow the search to proceed without improvement
CT_{max}	=	the starting temperature
CT_{min}	=	the finishing temperature

4.4 Constraint Violation Reduction Phase

The Constraint Violation Reduction (CVR) phase is the main process for solving the constraint satisfaction problem (CSP) described in section 3.4.5. The heuristic in the CVR phase is developed based on the tabu search approach. In this section, a knowledge-based system that provides constraint violation information to help facilitate the search process is defined. The basic idea is that information from the current violated constraints is used to navigate the search space and to change the values of potential variables that can contribute to the reduction of constraint violations.

The following provides an overview of the tabu search algorithm.

4.4.1 Overview of Tabu Search

Tabu search (TS) was first suggested by Glover [84]. The basic idea of TS is for the solution space to be explored economically and efficiently using memory strategies to guide the search process. Webster's II dictionary defines tabu (or taboo) as "A prohibition against doing, using, or mentioning something because of association with powerful supernatural forces" [85]. In this application, TS imposes restrictions on the available choices in order to guide the search process and avoid cycling of solutions during particular iterations, thereby escaping from local optima. Restrictions are imposed or created by making reference to memory structures which are designed for this specific purpose.

The main elements of the TS algorithms are:

- Evaluation function: An evaluation function maps a solution into a value representing the cost of the solution to the problem considered.
- Neighborhood and move operator: In each iteration, the TS algorithm selects a new solution which is the best unrestricted solution in the neighborhood of the current solution. The neighbor solutions are generated by a perturbing function called a move operator, which specifies attribute(s) of the current solution to be adjusted. All parameters of the new solution are the same as those of the current solution except for the parameter(s) that are modified by the move operator.
- Memory: The memory used in tabu search is both explicit and attributive. Explicit memory records complete solutions (called elite solutions) visited during the search. The memorized elite solutions are used to guide the search to revisit a solution area more thoroughly or to direct the search to areas that have been visited rarely.

Alternatively, TS uses attributive memory which records information about solution attributes that change when a new solution is generated from the current one. For example, in a routing problem, attributes can consist of nodes or links that are added or dropped by a move operator. The main memory mechanisms include short term memory and long term memory.

- Short term memory: The short term memory commonly keeps track of solution attributes changed in recent iterations. The changed attributes are labeled tabu-active, and their status remains active for a number of iterations called the tabu tenure. There are two types of tabu tenure: static and dynamic tabu tenure. In static tabu tenure, the tabu active status is maintained for a fixed number of iterations. In dynamic tabu tenure, the number of iterations for which the attributes remain active is variable. Solutions that contain tabu-active elements or combinations of tabu-active attributes become forbidden and are not included in the neighborhood. The tabu (restricted) solution can be selected as the new solution only if it satisfies the aspiration criteria explained below.
- Long term memory: Long term memory makes use of the attributive and explicit memory mechanisms. In the attributive-based long term memory, the mechanism keeps track of how often attributes change and whether the changes lead to good solutions or not. In the explicit-based long term memory, the mechanism keeps track of elite solutions found during the search. Later, the search can revisit the area of the elite solutions more thoroughly.
- Aspiration criteria: Normally, attributes that are tabu-active are not included in the neighborhood of the current solution. However, the tabu-active status can be

overridden when certain condition(s), called aspiration criteria, are met. An example of aspiration criteria is when the tabu-active solution is better than any other solution seen thus far in the search. Such a condition is called the improved-best aspiration criterion [84].

- Termination rule: The termination rule specifies when to stop the TS search.

The following describes the main components of the tabu search methods, including the overall tabu search structure of the CVR phase, the evaluation function, the structure of the constraint violation information, the neighborhood structure and move operators, the short term memory, the long term memory, the aspiration criteria, and the termination rules.

4.4.2 Overall Structure of Tabu Search Procedures Implemented for the CVR Phase

The CVR phase begins by evaluating the initial network configuration (initial solution) generated during the first two phases (the construction and FCA phases). If any design requirement is violated, the CVR process evaluates the neighbor solutions of the initial solution and moves to the best unrestricted solution (the solution with the least degree of constraint violation) in the neighborhood. Then, the best unrestricted solution is regarded as the current solution during the following iteration and is evaluated. The process of neighborhood evaluation is repeated until either a feasible solution satisfying all design constraints is found or the termination rule is reached. Evaluating all solutions in the neighborhood can take a lot of computational time because each evaluation requires the computation of the received signal strength and the SIR levels at all signal test points (STPs). The number of STPs depends on the size of the service area and the grid spacing used. Moreover, the size of the neighborhood can be large, depending on the size of the domain of each variable. To restrict the number of neighbor solutions examined in each iteration, we utilize the constraint violation information to derive a set of potential neighbor solutions that are likely to lead the search to a feasible solution.

Figure 4.7 shows the structure of the tabu search procedure implemented during the CVR phase. The process begins with an evaluation of an input WLAN configuration (denoted as the current solution s) to determine if it satisfies all design requirements. If not, the violated constraints are determined and potential neighbor solutions are derived based on the neighborhood structure and move operators described in section 4.4.5. Let $\mathbf{S} = \{s_1, s_2, s_3, \ldots, s_m\}$ be a set of neighbor solutions generated based on the current solution s.

To prevent a circulation in the search, the short term memory of the tabu search technique is applied. The attributes of solutions that are recorded during the search include direction and power level attributes. For each candidate solution $s_i \in \mathbf{S}$, attributes of s that are changed to create s_i are identified. For example, if the location of AP1 is changed to create s_1, then the direction that AP1 is moved is computed. In this example, the attribute of s that is changed to create s_1 is the direction attribute.

The tabu status of the changed attributes is checked. If the tabu status is not active, the solution $s_i \in \mathbf{S}$ is then evaluated; if it is better than other candidate solutions in $\mathbf{S}$, it is recorded as the best solution in $\mathbf{S}$ (denoted as s^*). If the tabu status of some attributes is active, the solution s_i is then evaluated while the tabu-active attributes remain unchanged; if it is better than other candidate solutions in $\mathbf{S}$, it is recorded as the best in $\mathbf{S}$.

The aspiration criteria are applied to determine when the tabu activation rules can be overridden. The improved-best aspiration criterion is applied here by allowing an evaluation of s_i whose tabu-active attributes are changed. If it is better than both the best_so_far solutions (denoted as s^{**}) and the best solution in **S** (denoted as s^*), s^{**} and s^* will be updated.

After all s_i in **S** are evaluated in each iteration, the best solution in the neighborhood s^* is identified. The search moves to s^*. In other words, in the next iteration, the current solution s is set to s^*. The iterative process of the CVR phase stops when either a feasible configuration is found or the CVR stopping rule is met.

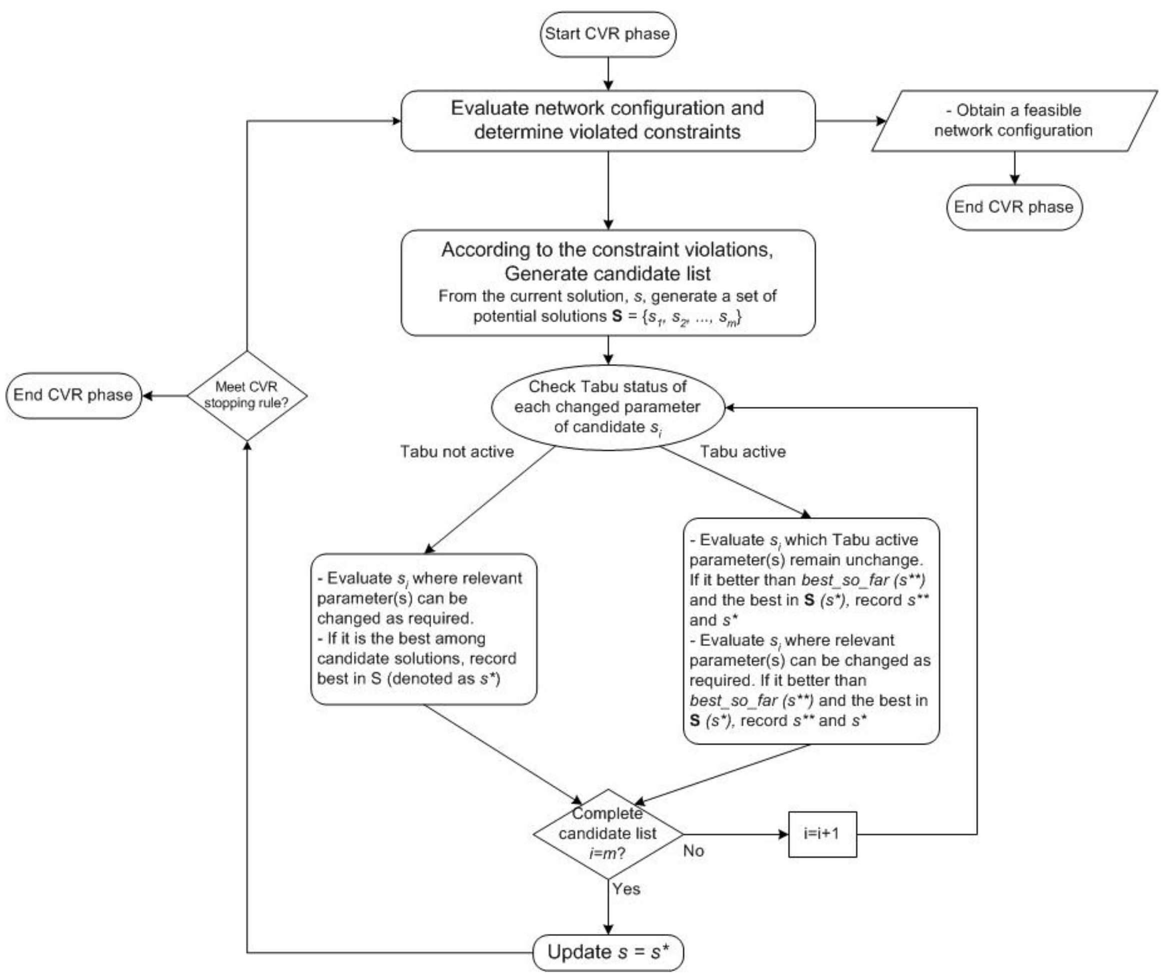


Figure 4.7 Structure of Tabu search procedure implemented for the CVR phase

4.4.3 Evaluation of Network Configuration

The evaluation function implemented here is a measure of the degree of constraint violation of the network configuration. It is equal to the sum of two weighted functions, each of which represents a degree of constraint violations for each design criterion. The first function β_1 measures the violations of the radio signal coverage criterion. It is equal to the sum of the normalized signal strength violation and normalized SIR violation at Signal Test Points (STPs). This normalized sum value is weighted with a factor (w_g) representing the relative importance of each STP. Note that the w_g can be different across STPs in the service area if imposing a different priority to different locations in the service region is desired. When all usable locations have the same service priority, w_g is set to 1. The total violation of the first design criterion across all STPs is then normalized by $2\sum_{\forall g \in G} w_g$ so that the value of β_1 is scaled from 0 to 1.

The second function, β_2, measures violations of the data rate requirement criterion. It is equal to the sum of the normalized data rate violation for each wireless user. The total violation of the second design criterion across all active wireless users is then normalized by $\sum_{\forall i \in M} w_i$ so that the value of β_2 is scaled from 0 to 1. In the evaluation function, the weight factors w_1 and w_2 are used to impose different focuses in the design. If the desire is to provide only radio signal coverage in the service area, the weight factor can be set as $w_1 = 1$ and $w_2 = 0$. If both design criteria are required, one can set $w_1 = 1$ and $w_2 = 1$.

$$E(solution_i) = w_1\beta_1(\text{signal coverage criteria}) + w_2\beta_2(\text{demand coverage criteria}) \qquad (4.9)$$

where

w_1 = weight factor representing relative importance of radio signal coverage criteria

w_2 = weight factor representing relative importance of traffic demand coverage criteria

$$\beta_1(.) = \frac{1}{2\sum_{\forall g\in G} w_g} \sum_{\forall g\in G} w_g \left(max\left(0, \frac{Pr_{threshold} - Pr_g}{Pr_{threshold}}\right) + max\left(0, \frac{SIR_{threshold} - \left(Pr_g / Intf_g\right)}{SIR_{threshold}}\right)\right)$$

(4.10)

$$\beta_2(.) = \frac{1}{\sum_{\forall i\in\Lambda} w_i} \sum_{\forall i\in\Lambda} w_i \left(max\left(0, \frac{d_t - r_i^t}{d_t}\right)\right) \qquad (4.11)$$

where

$Pr_{threshold}$ = receiver sensitivity threshold (in watt)

Pr_g = signal strength at STP g (in watt)

$Intf_g$ = interference level at STP g (in watt)

$SIR_{threshold}$ = signal to interference ratio threshold (no unit, ratio)

d_t = required average data rate of user of type t (in bps)

r_i^t = achievable average data rate of user i which is a user of type t (in bps)

w_g = weight factor representing relative importance of grid g

w_i = weight factor representing relative importance of user i

4.4.4 Constraint Violation Information

The direction of the search is determined by exploiting data structures of the information regarding constraint violations. In each iteration, the current solution might violate some constraints. In other words, signal strength and/or SIR level in some areas might be below the threshold and/or some APs might experience a traffic overload. The idea is to keep track of those location(s) where network performance does not meet the design requirements. This information is used to identify variables that should be adjusted to improve network performance based on the specified design criteria.

To manage the information about constraint violations, we divide the service area into a unit area (UA) where the performance information described below is evaluated and recorded. The reason that the constraint violation information is managed per UA is to provide a good indicator about the locations of problems, in case those locations are spread across the service region. An example described later in this section will illustrate this scenario.

The size of the UA is a parameter to be tuned. Chapter 5 illustrates experiments conducted to define an efficient UA size.

Notations:

n_{UA} = Total number of UAs managed in the service area

G_i^{pr} = Set of STPs in UA_i where the received signal strength is below the threshold, for $i = 1,\ldots, n_{UA}$

G_i^{SIR} = Set of STPs in UA_i where the SIR level is below the threshold, for $i = 1,\ldots, n_{UA}$

A_i^{ov} = Set of APs in UA_i that cannot provide the required average data rate to associated wireless users for $i = 1,\ldots, n_{UA}$

For each UA_i, the following information is observed and recorded.

1. Information related to low signal strength grids

- Center of gravity of STPs in G_i^{pr} (denoted by CoG_i^{pr})
- The lowest received signal strength among STPs in G_i^{pr} (denoted by l_i^{pr})
- Location of the STP with the lowest signal strength (denoted by Loc_i^{pr})

2. Information related to low SIR grids

- $\left|G_i^{SIR}\right|$, Number of STPs in G_i^{SIR}
- Average SIR level of all STPs in G_i^{SIR} (denoted by avg_i^{SIR})
- Center of gravity of STPs in G_i^{pr} (denoted by CoG_i^{SIR})

3. Information related to overload APs

- A_i^{ov}
- Center of gravity of the locations of wireless users assigned to $a^{ov} \in A_i^{ov}$ (denoted by $CoG_i^{a^{ov}}$) where a^{ov} denotes an overload AP

Figure 4.8 illustrates the constraint violation information observed within the example service area. The red grids represent the STPs where signal strength is below the received sensitivity threshold. The red cross-circle represents the centers of gravity for those grids with low signal strength. The blue grids represent the STPs where SIR is below the acceptable threshold. The blue cross-circle represents the center of gravity for those grids with low SIR level. The small green circles represent those wireless users whose average data rate is below the required rate. The green cross-circle represents the center of gravity for those wireless users.

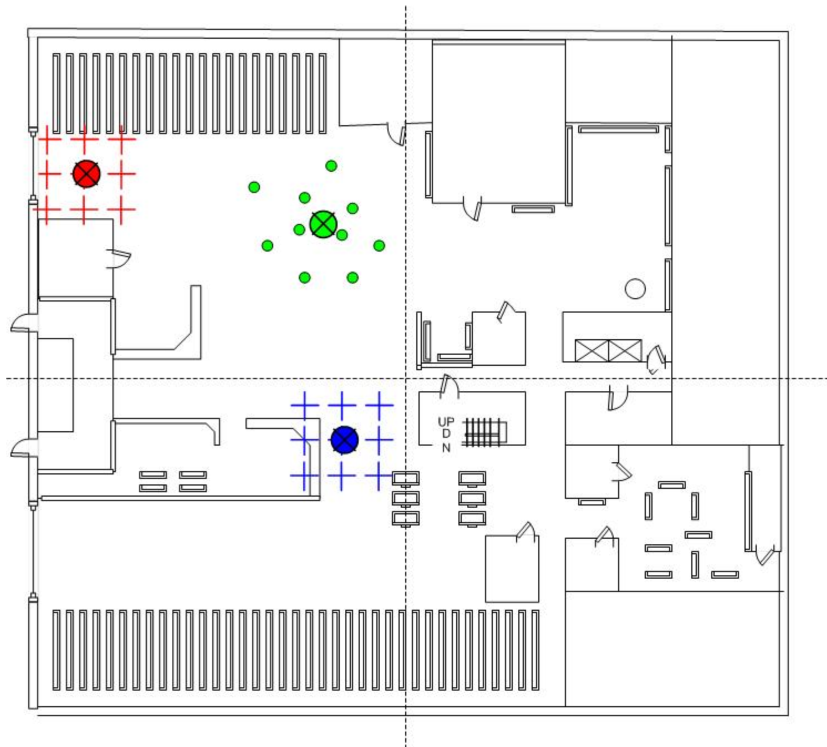


Figure 4.8 Examples of constraint violation information management

4.4.5 Neighborhood Structure and Move Operators

The neighbors of the current solution are obtained by adjusting parameters of APs which are likely to result in the reduction of constraint violations. The constraint violation information described in section 4.4.4 is used to guide the derivation of a set of potential neighbor solutions. This information is used to identify a set of potential APs denoted as $\boldsymbol{A_{potential}}$ whose variables should be adjusted to improve the network performance. We describe how to determine $\boldsymbol{A_{potential}}$ later in this section.

To generate neighbors of the current solution, two types of move operators are developed. They include greedy moves and conservative moves. Each type deals with changing the power level (p) and the location (x, y) of APs in the set $\boldsymbol{A_{potential}}$.

In greedy moves, parameters (x,y and/or p) of more than one AP in the set $\boldsymbol{A_{potential}}$ are changed simultaneously. The advantage of this strategy is that the constraint violations in multiple areas can be improved in one iteration. This allows a fast convergence of the search. However, greedy moves can be too aggressive and cause unintended constraint violation in some cases, especially when the search is approaching a feasible solution and only a small violation remains. Conservative moves overcome this problem by changing only parameter (x,y) or parameter p of one AP in the set $\boldsymbol{A_{potential}}$ at a time. Another advantage of the conservative move strategy is that it allows the search to escape from those solutions in which the greedy moves could not derive any better solution. A detailed description of greedy and conservative moves is presented later in this section.

The parameter frequency channel (f) is not changed by either greedy moves or conservative moves. The reason is that each type of move intends to reduce the constraint violations by adjusting APs' locations and power levels based on the current setting of frequency assignment. Independently changing the frequency channel of some APs could increase interference in multiple areas around the AP(s). However, in each iteration considering the locations and power levels of APs of the current solution, a separate process of frequency channel assignment is performed by the simulated annealing method described in section 4.3. The obtained frequency assignment is considered as another neighbor solution of the current solution, while the locations and power level of APs remain unchanged.

According to the type of constraint violations and the APs' parameters that need to be adjusted, the neighborhood structures are classified into four types as follows:

a. Neighborhood structure of the current solution that causes signal strength violation
b. Neighborhood structure of the current solution that causes interference violation
c. Neighborhood structure of the current solution that causes average data rate violation
d. Neighborhood structure of the current solution that requires frequency channel reassignment

The following describes how to identify $\boldsymbol{A_{potential}}$ and explains in detail the greedy and conservative moves implemented for each neighborhood structure.

Notations:

$A^{ssv}_{potential}$	=	Set of potential APs selected with respect to the signal strength violation (*ssv*)
$A^{SIRv}_{potential}$	=	Set of potential APs selected with respect to the SIR violation (*SIRv*)
$A^{drv}_{potential}$	=	Set of potential APs selected with respect to the average data rate violation (*drv*)
$A_{potential}$	=	Set of all potential APs = $A^{ssv}_{potential} \cup A^{SIRv}_{potential} \cup A^{drv}_{potential}$
$\Omega^{ssv}_{g.m.}$	=	Set of neighbor solutions regarding the signal strength violation derived by the greedy moves (g.m.)

$\Omega^{ssv}_{c.m.}$ = Set of neighbor solutions regarding the signal strength violation derived by conservative moves (c.m.)

$\Omega^{SIRv}_{g.m.}$ = Set of neighbor solutions regarding the SIR violation derived by greedy moves

$\Omega^{SIRv}_{c.m.}$ = Set of neighbor solutions regarding the SIR violation derived by conservative moves

$\Omega^{drv}_{g.m.}$ = Set of neighbor solutions regarding the average data rate violation derived by greedy moves

$\Omega^{drv}_{c.m.}$ = Set of neighbor solutions regarding the average data rate violation derived by conservative moves

a. Neighborhood structure of the current solution that causes signal strength violation

Step 1: Identify $A^{ssv}_{potential}$

For each UA_i, if there exist STPs whose received signal strength is below the received sensitivity threshold, the AP closest to CoG^{pr}_i is selected as a potential AP for UA_i denoted as a^{pr}_i. After evaluating signal strength constraint violation in all UAs, we obtain $A^{ssv}_{potential}$.

$$A^{ssv}_{potential} = \{ a^{pr}_i \mid \left|G^{pr}_i\right| > 0, \forall i = 1,\ldots,n_{UA} \}$$

Step 2: Generate $\Omega^{ssv}_{g.m.} = \{ s_1, s_2, s_3 \}$

To generate each solution in this neighborhood, the greedy moves intend to adjust parameters ((*x*,*y*) and/or *p*) of all APs in $A^{ssv}_{potential}$ simultaneously with a goal to improve the radio signal strength in multiple areas (UAs) in one iteration. The following describes the greedy move operation to generate $\Omega^{ssv}_{g.m.}$.

s_1: The first solution is obtained by changing locations of all APs in $A^{ssv}_{potential}$. The new location of each AP in $A^{ssv}_{potential}$ is determined by a line search heuristic which works as follows. Move a_i^{pr} toward CoG_i^{pr} by a fixed step size (λ) at a time until it provides sufficient signal strength to the Loc_i^{pr}. It then stops and the new location of a_i^{pr} is obtained. Repeat this process for all a_i^{pr} in $A^{ssv}_{potential}$. Finally, a solution that provides new locations to all APs in $A^{ssv}_{potential}$ is obtained.

s_2: The second solution is obtained by increasing the transmission power of all APs in $A^{ssv}_{potential}$ by one level, providing that the power level does not exceed the maximum power level.

s_3: The third solution is obtained by a combination of adjusting locations and power levels of APs in $A^{ssv}_{potential}$. The power level of the AP that is selected as a potential AP for more than one UA is increased by one level. The location of the AP that is selected as a potential AP for only one UA will be adjusted, an the new location will be derived from the line search method described above.

Step 3: Generate $\Omega^{ssv}_{c.m.}$

Each solution in this neighborhood is obtained by the conservative moves that attempt to change only one parameter (location (x,y) or power level, p) of one AP in $A^{ssv}_{potential}$. We consider eight directions to move each potential AP as illustrated in Figure 4.9. For each a^{pr}_i in $A^{ssv}_{potential}$, the first neighbor solution is obtained by changing a^{pr}_i's location to the direction 1 by a step size λ. The second solution in the neighborhood is obtained by changing a^{pr}_i's location to the direction 2 by a step size λ and so on for all eight directions. This operation is repeated for all APs in $A^{ssv}_{potential}$. We also consider increasing the power level of each potential AP by one level. Thus, the total number of solutions in the $\Omega^{ssv}_{c.m.}$ neighborhood is equal to $9 \times \left|A^{ssv}_{potential}\right|$.

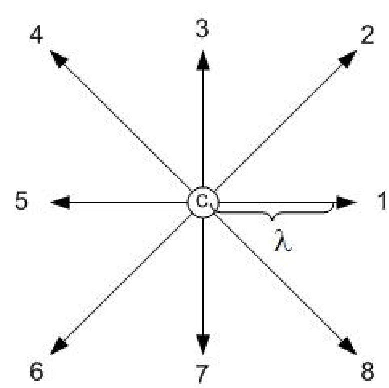


Figure 4.9 Direction to move AP for conservative move operations

b. Neighborhood structure of the current solution that causes interference violation

Step 1: Identify $A_{potential}^{SIRv}$

The STPs in each UA may obtain their highest SIR levels from different APs located in nearby locations. Figure 4.10 illustrates this scenario. AP1 and AP4, located in UA1 and UA4, respectively, operate at frequency channel 1 while AP2 and AP3, located in UA2 and UA3, operate at frequency channels 6 and 11, respectively. Signals from AP1 and AP4 can interfere with each other at the boundary of the AP coverage area. So, those STPs in the boundary of the coverage area of AP1 and AP4 may obtain the highest SIR level from AP2 and AP3. In Figure 4.10, green grids represent STPs in UA1 that obtain the highest SIR from AP1 while red and blue grids represent STPs in UA1 that obtain the highest SIR from AP2 and AP3, respectively.

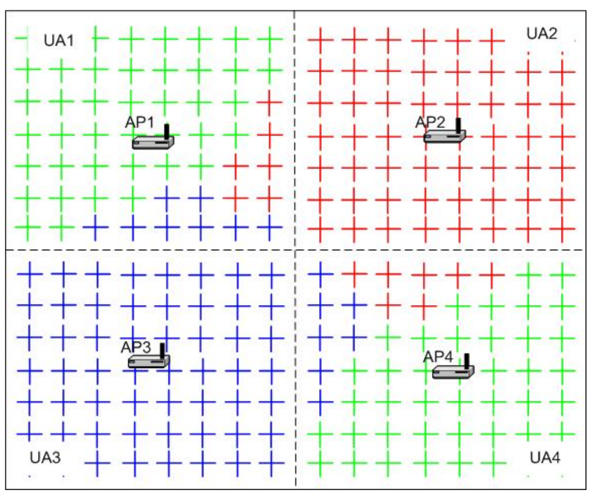


Figure 4.10 Example where STPs in a particular UA obtain their highest SIR levels from different APs

Based on the concept depicted in Figure 4.10, we adopt the following notation:

G_i^{SIR} = Set of STPs in UA_i where the SIR level is below the $SIR_{threshold}$, for i = 1,…, n_{UA}

G_i^{SIR} is subdivided into σ subsets where

$G_{i,1}^{SIR} \cup G_{i,2}^{SIR} \cup \ldots \cup G_{i,k}^{SIR} \cup \ldots \cup G_{i,\sigma}^{SIR} = G_i^{SIR}$ and

$G_{i,1}^{SIR} \cap G_{i,2}^{SIR} \cap \ldots \cap G_{i,k}^{SIR} \cap \ldots \cap G_{i,\sigma}^{SIR} = \varnothing$

σ = Total number of APs that provide the highest SIR level to STPs in G_i^{SIR}

$a_{i,k}^h$ = AP that provides the highest SIR level to STPs in $G_{i,k}^{SIR}$, for k = 1, …, σ

$a_{i,k}^{INTF}$ = AP that is closest to $a_{i,k}^h$ and operates at the interfering frequency channel with respect to the frequency channel assigned to the $a_{i,k}^h$. Note that in the case of 802.11b, the interfering frequency channel is the channel where the channel distance to the frequency channel of the $a_{i,k}^h$ is less than five. In the case of the 802.11a, the AP whose frequency channel is the same as that of $a_{i,k}^h$ is the interfering AP.

$A_{potential}^{SIR,h}$ = Set of potential APs that could improve the received signal strength so that the SIR level can be improved

$A_{potential}^{SIR,INTF}$ = Set of potential APs that could reduce the interference signal so that the SIR level can be improved

For each UA_i, if there are STPs whose SIR level is below the SIR threshold, among all k = 1, …, σ, denote the $a_{i,k}^h$ of the set $G_{i,k}^{SIR}$ that contains the highest number of

STPs as $a_{i,k}^{h,\max}$ and compute the center of gravity of STPs in such $G_{i,k}^{SIR}$ (denoted by $CoG_{i,k}^{SIR}$). Also determine $a_{i,k}^{INTF}$ that corresponds to $a_{i,k}^{h,\max}$ (denoted as $a_{i,k}^{INTF,\max}$). After evaluating the constraint violation regarding SIR violation in all UAs, $A_{potential}^{SIR,h}$ is obtained:

$A_{potential}^{SIR,h}$ = a set of $a_{i,k}^{h,\max}$ for UA_i that $\left|G_i^{SIR}\right| > 0$ for all $i = 1, \ldots, n_{UA}$ and $G_{i,k}^{SIR}$ is the set that contains the highest number of STPs where $1 < k < \sigma$

$A_{potential}^{SIR,INTF}$ = a set of $a_{i,k}^{INTF}$ that corresponds to all $a_{i,k}^{h,\max}$ in $A_{potential}^{SIR,h}$

Then $A_{potential}^{SIRv}$ is obtained:

$$A_{potential}^{SIRv} = A_{potential}^{SIR,h} \cup A_{potential}^{SIR,INTF}$$

Step 2: Generate $\Omega_{g.m.}^{SIRv} = \{ s_1, s_2, s_3, s_4, s_5, s_6 \}$

To generate each solution in this neighborhood, the greedy moves intend to adjust location and/or power level ((x,y) and/or p) of all APs in $A_{potential}^{SIRv}$ in one move with the goal of improving the SIR level in multiple areas (UAs) in one iteration. The following describes the greedy move operation to generate $\Omega_{g.m.}^{SIRv}$.

s_1: The first solution is obtained by moving all APs in $A_{potential}^{SIR,INTF}$ away from their corresponding $CoG_{i,k}^{SIR}$ in order to reduce the interference signal of those STPs in the set $G_{i,k}^{SIR}$. The new location (x'_j, y'_j) of each $a_j \in A_{potential}^{SIR,INTF}$ is heuristically obtained as follows:

$$(x'_j, y'_j) = (x_j + \Delta x'_j, y_j + \Delta y'_j) \quad (4.12)$$

where

$(\Delta x_j, \Delta y_j)$ is the step size to move the $a_j \in A_{potential}^{SIR,INTF}$ and is computed by

$$\Delta x_j = \left(\frac{SIR_{threshold} - SIR_{avg}}{SIR_{threshold}} \right) . \left| x_j - x\_CoG_{i,k}^{SIR} \right| \tag{4.13}$$

$$\Delta y_j = \left(\frac{SIR_{threshold} - SIR_{avg}}{SIR_{threshold}} \right) . \left| y_j - y\_CoG_{i,k}^{SIR} \right| \tag{4.14}$$

$$\Delta x'_j = \oplus \min\left(upperbound , \max\left(\Delta x_j , lowerbound \right) \right) \tag{4.15}$$

$$\Delta y'_j = \oplus \min\left(upperbound , \max\left(\Delta y_j , lowerbound \right) \right) \tag{4.16}$$

where

$(x_j, y_j) = (x, y)$ coordinate of $a_j \in A_{potential}^{SIR,INTF}$

$(x\_CoG_{i,k}^{SIR}, y\_CoG_{i,k}^{SIR}) = (x, y)$ coordinate of the center of gravity $CoG_{i,k}^{SIR}$

$\oplus$ represents the direction to move $a_j \in A_{potential}^{SIR,INTF}$. (It indicates whether to move away from or toward $CoG_{i,k}^{SIR}$)

lowerbound specifies the minimum step size to move AP

upperbound specifies the maximum step size to move AP

Repeat this process for all $a_j \in A_{potential}^{SIR,INTF}$ to obtain a neighbor solution s_1 whose parameter (x_j, y_j) of all $a_j \in A_{potential}^{SIR,INTF}$ is moved to new location (x'_j, y'_j).

s_2: The second solution is obtained by moving all APs in $A_{potential}^{SIR,h}$ toward their corresponding $CoG_{i,k}^{SIR}$ in order to increase the signal strength level so that the SIR level can be improved. The new location of each $a_j \in A_{potential}^{SIR,h}$ can be obtained by the same method described in equations (4.12) – (4.16).

s_3: The third solution is obtained by reducing the power level of all APs in $A_{potential}^{SIR,INTF}$ by one level, providing that the resulting power level is not lower than the minimum power level. The reduction of power level is intended to lower the interference signal.

s_4: The fourth solution is obtained by adjusting a combination of locations and power levels of APs in $A_{potential}^{SIR,INTF}$. An AP which is selected as a potential AP ($a_{i,k}^{INTF,\max}$) for more than one UA can cause interference in multiple UAs. This AP's power level will be reduced by one level. The location of an AP that is a potential AP for only one UA will be adjusted by moving away from the corresponding $CoG_{i,k}^{SIR}$. The new location will be derived from the heuristic method described in equations (4.12) – (4.16).

s_5: The fifth solution is obtained by increasing the power level of all APs in $A_{potential}^{SIR,h}$ by one level, providing that the power level does not exceed the maximum power level. The increase of those APs' power levels is intended to raise the wanted signal strength level.

s_6: The sixth solution is obtained by adjusting a combination of locations and power levels of APs in $A_{potential}^{SIR,h}$. An AP which is selected as a potential AP ($a_{i,k}^{h,\max}$) for more than one UA can improve the wanted signal level in multiple UAs. Such AP's power level will be increased by one level. The location of the AP that is a potential AP for only one UA will be adjusted by moving toward the corresponding $CoG_{i,k}^{SIR}$. The new location will be derived from the heuristic method described in equations (4.12) – (4.16).

Step 3: Generate $\Omega_{c.m.}^{SIRv}$

Each solution in this neighborhood is obtained by conservative moves that attempt to change only one parameter (location (*x*,*y*) or power level, *p*) of one AP in $A_{potential}^{SIRv}$. Again we consider eight directions to move each potential AP as illustrated in Figure 4.9. For each AP in $A_{potential}^{SIRv}$, the first neighbor solution is obtained by changing the AP's location in direction 1 by a step size λ. The second solution in the neighborhood is obtained by changing the AP's location in direction 2 by a step size λ and so on for all eight directions. This operation is repeated for all APs in $A_{potential}^{SIRv}$. We also consider increasing the power level of each potential AP in $A_{potential}^{SIR,h}$ (one AP at a time) by one level and decreasing the power level of each potential AP in $A_{potential}^{SIR,INTF}$ (one AP at a time) by one level. Thus, the total number of solutions in this neighborhood is equal to

$$8 \times \left|A_{potential}^{SIRv}\right| + \left|A_{potential}^{SIR,h}\right| + \left|A_{potential}^{SIR,INTF}\right|$$

c. Neighborhood structure of the current solution that causes average data rate violation

Step 1: Identify $A_{potential}^{drv}$

For each UA_i, if there exist any overloaded AP, select a low-load AP (denoted as a_i^{low}) that is closest to the center of gravity of a set of wireless users associated with the overloaded AP ($CoG_i^{a^{ov}}$). If no low load AP exists in the nearby UAs, an additional AP is placed at the $CoG_i^{a^{ov}}$.

After evaluating for average data rate constraint violations in all UAs, we obtain $A_{potential}^{drv}$.

$$A_{potential}^{drv} = \text{a set of } a_i^{low} \text{ for } UA_i \text{ that } \left|A_i^{ov}\right| > 0 \text{ for all } i = 1, \ldots, n_{UA}$$

Step 2: Generate $\Omega_{g.m.}^{drv} = \{ s_1, s_2, s_3, s_4 \}$

s_1: The first solution is obtained by changing the locations of all APs in $A_{potential}^{drv}$. The new location of each $a_i^{low} \in A_{potential}^{rate}$ is determined by a line search heuristic which works as follows: Move a_i^{low} toward the corresponding $CoG_i^{a^{ov}}$ with a fixed step size (λ) at a time until it provides sufficient signal strength to the location of the corresponding $CoG_i^{a^{ov}}$ (i.e., until it can take some load from the overload AP). It then stops and the new location of a_i^{low} is obtained. Repeat this process for all selected APs in $A_{potential}^{rate}$. Finally, the solution that provides new locations to all APs in $A_{potential}^{rate}$ is obtained.

s_2: The second solution is obtained by increasing the power level of all low load APs in $A_{potential}^{rate}$ by one level, providing that the power level does not exceed the maximum power level. The increase of the APs' power level is intended to expand the coverage area of the low load APs so that they can take some load from the overloaded AP.

s_3*:* The third solution is obtained by decreasing the power level of the overloaded APs in A_i^{ov} by one level, providing that the power level is not lower than the minimum power level. The reduction of the APs' power level is intended to reduce the coverage area of the overloaded APs so that their traffic load can be reduced and a higher average user data rate can be obtained.

s_4*:* The forth solution is obtained by adjusting a combination of locations and power levels of APs in $A_{potential}^{drv}$. The power level of the AP that is selected as a potential AP (a_i^{low}) for more than one UA will be increased by one level. The location of the AP that is a potential AP for only one UA will be adjusted by moving it toward the corresponding $CoG_i^{a^{ov}}$. The new location is derived from the line search heuristic method described previously in s_1.

Step 3: Generate $\Omega_{c.m.}^{drv}$

Each solution in this neighborhood is obtained by conservative moves that attempt to change only one parameter (location (*x*,*y*) or power level, *p*) of one AP in $A_{potential}^{drv}$. We consider eight directions to move each potential AP as illustrated in Figure 4.9. For each AP in $A_{potential}^{drv}$, the first neighbor solution is obtained by changing the AP's location in direction 1 by a step size λ. The second solution in the neighborhood is obtained by changing its location in direction 2 by a step size λ and so on for all eight directions. Repeat this operation for all APs in $A_{potential}^{drv}$. We also consider increasing the power level of one AP$\in$ $A^{drv}_{potential}$ at a time to generate additional solutions for the $\Omega^{drv}_{c.m.}$ neighborhood. The total number of solutions in the $\Omega^{drv}_{c.m.}$ neighborhood is equal to $9\times\left|A^{drv}_{potential}\right|$

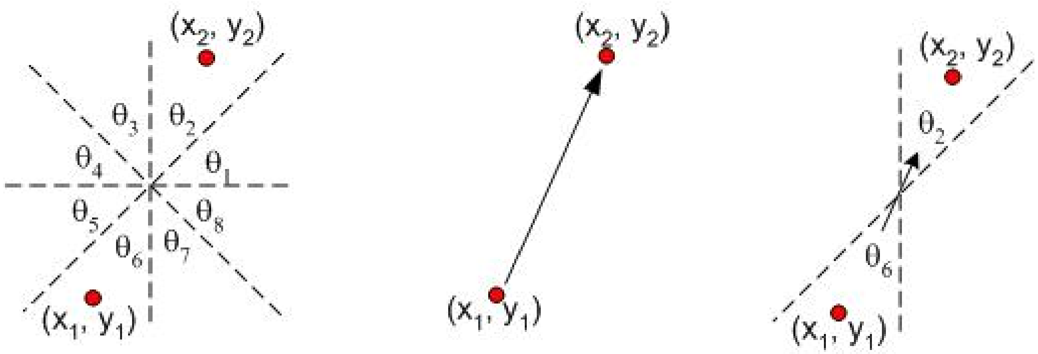
d. Neighborhood structure of the current solution that needs frequency channel reassignment

The APs' frequency channels are not changed in the neighborhood structures described previously. The reason is that we want to keep the APs' frequency assignments the same while adjusting their locations and power levels to improve the constraint violations. After those adjustments, the APs' frequency channels may need to be reassigned because the change in locations and power levels can change the number of channel separations required between AP pairs.

Here we consider the reassignment of the frequency channels while the APs' locations and power levels remain unchanged. The frequency channel reassignment is obtained by the simulated annealing method described in section 4.3. WLAN configuration with new frequency channel assignment is considered as another solution in the neighborhood of a particular WLAN configuration.

4.4.6 Short Term Memory

The memory structure implemented in the tabu search process of the CVR phase includes both short term memory and the long term memory. This section describes the short term memory. The algorithm keeps track of two attributes of the WLAN configuration, specifically the direction APs are moved and their power levels. For the direction (*di*) attribute, eight directions are managed (Figure 4.11a). Figure 4.11b illustrates an example in which an AP is moved from a location (x_1, y_1) to (x_2, y_2). In this example, the AP is moved from the θ_6 direction to the θ_2 direction. Thus, the tabu status of the θ_6 attribute is labeled active. This means that such an AP cannot be moved back to the θ_6 direction for a number of iterations called the tabu tenure. The direction attribute is maintained to prevent moving AP(s) back to the area where they were recently located.



a. Direction attributes b. Example of the direction attribute management

Figure 4.11 Short term memory: Direction attributes

For the power level (*pl*) attribute, increasing and decreasing-*pl* attribute are maintained. If an AP's power level is increased, the decreasing-*pl* attribute is labeled active. The tabu-active status of the decreasing-*pl* attribute of a particular AP means that its power level cannot be decreased during the tabu tenure unless it meets the aspiration criteria. The increasing-*pl* attribute works in the opposite manner. The use of these attributes can prevent the oscillating of an AP's power level.

4.4.7 Long Term Memory

For the long term memory, we employ the explicit memory mechanism, which records a complete solution (i.e., a detail network configuration), to keep track of good solutions observed along the search process. In each iteration of the CVR phase, the search selects and moves to the best unrestricted solution in the neighborhood with respect to the degree of constraint violations of the solution. The algorithm records the second best solution of the neighborhood in a list called the elite list. (As explained earlier, the solutions recorded in the elite list are called elite solutions.) The elite solutions are revisited in the intensification phase.

4.4.8 Aspiration criterion

An aspiration criterion is used to determine when tabu active status can be overridden. The improved-best aspiration criterion is implemented in the CVR procedure. The improved-best criterion allows the neighbor solution whose attribute direction and/or power level is active to be accepted as a next move if it is a better network configuration than the best solution found so far [84].

4.4.9 Stopping Criteria for CVR Phase

The CVR phase stops either when a feasible solution is found or when the number of consecutive iterations passed without improvement reaches a threshold value Max_{iter_CVR}. The latter case activates the intensification phase.

4.4.10 Static Parameters in the CVR Phase

Static parameters in the CVR phase that remain unchanged during the course of the CVR process include:

UA size	=	The size of the unit area that is used to partition the service area in order to manage the constraint violation information
λ	=	Step size to move AP
Max_{iter_CVR}	=	The number of consecutive iterations without improvement on the constraint violation reduction
$TabuTenure_{di}$	=	The tabu tenure of an AP's move direction attribute (the number of iterations during which the AP cannot return back to the direction from which it is moved)
$TabuTenure_{p}$	=	The tabu tenure of an AP's power level attribute (the number of iterations during which the AP's power level cannot be reduced (increased) if it has just been increased (reduced))
Elite_list_size	=	The number of good solutions recorded in the list

4.5 Intensification Phase

The intensification phase aims to revisit good solutions found during the earlier CVR phase and explore their neighborhoods more thoroughly. Figure 4.12 shows an outline of the intensification phase. When the termination rule of the CVR phase is met, the search proceeds to the intensification phase where good solutions recorded in the elite list are revisited. The search restarts with the first solution in the elite list and the CVR process is performed until either a feasible solution is found or the termination rule of the CVR phase is met. In the latter case, the intensification phase proceeds to check the next solution in the elite list. After exploring all solutions in the record, if no feasible solution is found, additional AP(s) will be added.

```
Begin intensification phase
do {
        Choose solution from the elite list
        Perform CVR process and add new solutions to elite list when applicable
} while( Elite list not empty)
```

Figure 4.12 Outline of the intensification phase

4.6 Add-AP Phase

After the CVR and the intensification phases, if a feasible solution still has not been found, additional AP(s) will be added to the service area. Figure 4.13 depicts the procedure for adding an AP. If coverage holes or overloaded AP(s) exist, the unit area

(UA) that has the highest number of STPs with signal strength below the threshold and the AP with the highest traffic overload are defined. Next, the center of gravity where an additional AP will be placed is then defined. The new AP is initiated at the lowest power level. The frequency channel assignment is performed by the procedures described in the FCA phase. Finally, the new network configuration is input to the CVR phase.

```
If coverage hole exists
{
    Define UA with the highest number of STPs whose signal strength is below threshold
    Place an additional AP at the center of gravity of such STPs
}
Else if an overloaded AP exists
{
    Define an AP with the highest traffic overload
    Place an additional AP at the center of gravity of the demand nodes associated with
      the overloaded AP
}
Initiate the power level of the new AP at the lowest level
Perform frequency channel assignment
Return to CVR phase
```

Figure 4.13 Add-AP procedure

4.7 Summary

This chapter describes a heuristic solution technique developed to solve the demand-based CSP WLAN design problem. The heuristic solution technique aims to determine a feasible network configuration that can satisfy both the radio signal coverage requirement and the data rate capacity requirement. In particular, the heuristic solution technique produces a network configuration that specifies the number of APs required and their parameters, including locations, power levels and frequency channels. The developed heuristic solution technique consists of five phases. The first two phases, called the construction phase and the Frequency Channel Assignment (FCA) phase, aim to generate a good starting network configuration and provide an estimated number of APs required and their parameters. The third phase, called the Constraint Violation Reduction (CVR) phase, utilizes the constraint violation information to guide the search to converge to a feasible network configuration. The fourth phase, called the intensification phase, is performed only if no feasible solution is found during the CVR phase. This phase revisits good configurations found during the CVR phase and explores their neighborhoods more thoroughly. Finally, the fifth phase, called the add-AP phase, considers adding additional APs to the service area if the search has been unable to find a feasible network configuration. Parameters in the heuristic solution technique are defined through extensive numerical experiments presented in the next chapter.

5.0 SENSITIVITY ANALYSIS AND EFFECTIVENESS OF THE HEURISTIC SOLUTION TECHNIQUE

Here we present a sensitivity analysis of the key parameters that control operations of the proposed solution technique to the demand-based WLAN CSP. Through extensive numerical experiments, appropriate values of the key parameters that will be used in experimental studies discussed in chapter six are herein defined. This chapter is organized into four sections. Section 5.1 discusses parameters in demand-based WLAN design. Section 5.2 and 5.3 present a sensitivity analysis of the key parameters in the solution technique developed to solve the demand-based WLAN design problem. Section 5.4 presents an evaluation of the effectiveness of the proposed heuristic solution technique.

5.1 Parameters in the Demand-based WLAN Design Methodology

Parameters associated with the demand-based WLAN design can be categorized into two groups. The first group includes those parameters associated with characteristics of service scenarios, such as radio propagation, user distribution and traffic profile. These parameters are discussed in chapter six of this dissertation. The second group includes those parameters that control the operation of the heuristic solution technique, such as the static parameters in the Frequency Channel Assignment (FCA) phase and the static parameters in the Constraint Violation Reduction (CVR) phase. A sensitivity analysis of these parameters is presented in this chapter (sections 5.2 and 5.3, respectively).

5.2 Sensitivity Analysis of Static Parameters in the FCA Phase

The frequency channel assignment (FCA) phase relies upon the simulated annealing (SA) approach which is described earlier in this dissertation (section 4.3). The static parameters involved in the FCA phase include the initial temperature (CT_{max}), the parameter φ that specifies both how the temperature is reduced and the number of iterations at each temperature is increased, and the stopping criteria which is the finishing temperature (CT_{min}) and the number of consecutive iterations (*Max_noimprov*) that allow the search to proceed without improvement. The remainder of section 5.2 presents a sensitivity analysis of these parameters and determines the appropriate values to be used in subsequent numerical experiments.

5.2.1 Parameter Settings for the Sensitivity Analysis

The following describes the parameter settings for the sensitivity analysis involving CT_{max}, φ, CT_{min} and *Max_noimprov*:

- CT_{max} is an initial temperature that determines the probability of accepting an inferior FCA solution[†] at the beginning of the simulated annealing process. Recall that the acceptance probability (P_{accept}) for the minimization problem is written as

$$P_{accept} = e^{-\frac{\Delta \cos t}{CT}} \tag{5.1}$$

[†] Based on the cost function written in equation (4.5), an inferior FCA solution is a solution that has higher cost when compared to the initial FCA solution.

where $\Delta cost$ is the difference between the cost of a neighbor FCA solution and the cost of the current FCA solution. At the first temperature step, $CT = CT_{max}$ and the initial FCA solution is considered to be the current solution. CT_{max} is chosen so that during the iterations at this temperature step, the probability of accepting a worse FCA solution is approximately equal to a predefined initial P_{accept} (called P_{a_init}). Thus, CT_{max} is a function of $\Delta cost$ and P_{a_init}. It can be computed from

$$CT_{max} = -\frac{\Delta \cos t}{ln\left(P_{a\_init}\right)} \tag{5.2}$$

Kirkpatrick [86] argues that the value of P_{a_init} has to be high enough to allow a high probability of accepting inferior FCA solution in early iterations. He suggests that P_{a_init} be given a value of 0.8 [86]. In this study, we examine three levels of P_{a_init} : 0.3, 0.5, and 0.8. Each value's impact on the efficiency of the FCA process is observed.

The $\Delta cost$ used in Equation (5.2) is calculated in several steps. First, a number of random FCA solutions are generated and the $\Delta cost$ of all cost increasing FCA solutions as compared to the initial FCA solution (i.e., $\Delta cost > 0$) is recorded. Then, an average of those $\Delta cost$ values is computed. Finally, CT_{max} is obtained by applying the resulting $\Delta cost$ and the P_{a_init} under consideration in equation (5.2).

- For the parameter φ specifying how the temperature is reduced and how the number of iterations at each temperature is increased, we test four levels of φ: 0.3, 0.5, 0.7, and 0.9 to examine each value's effects on the simulated annealing process in the frequency channel assignment phase.
- Stopping rules for the FCA phase are specified by the CT_{min} or *Max_noimprov* parameter, whichever is reached first. CT_{min} specifies a minimum temperature level for the simulated annealing process while *Max_noimprov* specifies the number of consecutive

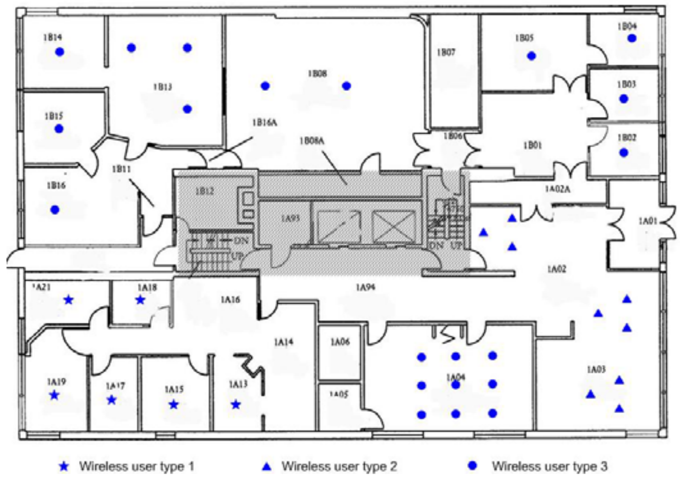
iterations that can be executed without improvement of the obtained solution. In this sensitivity analysis study, we test several values of $Max\_noimprov$ for each setting of φ and P_{a_init} as listed in Table 5.2 and 5.3. The value of CT_{min} is set at 0.01 for all cases.

5.2.2 Test Networks

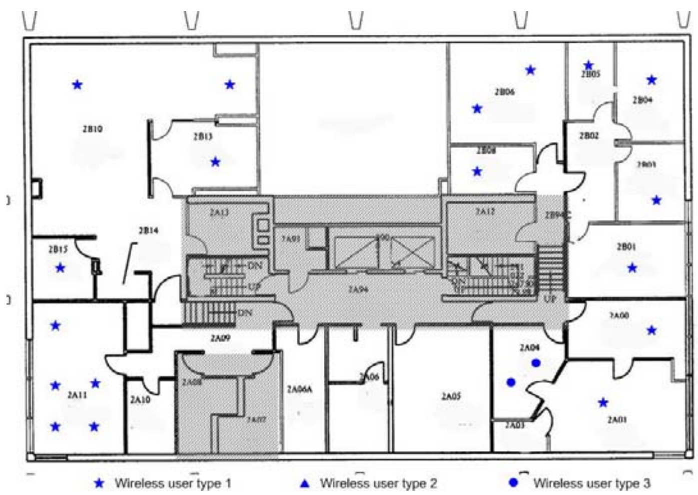
In order to study the effects of the key parameters described above, we conducted frequency channel assignment experiments on test networks in the School of Information Science (SIS) building at the University of Pittsburgh. The SIS building is an eight-story campus building comprised of classrooms, offices, laboratories, student lounges, and a library. The floor plan of the SIS building with expected wireless users indicated is presented in Figure 5.1. Network usage characteristics are summarized in Table 5.1. We consider two test network scenarios. The first test network (Network 1) considers the service area ranging from the first floor through the fourth floor of the building. This scenario requires 10 APs for network service. The second test network (Network 2) considers the service area ranging from the first floor through the eighth floor and requires 19 APs.

The sensitivity analysis detailed in this section focuses on the efficiency of the frequency channel assignment process of the simulated annealing method. The performance measures of this study includes the CPU time to run the simulated annealing process, the value of the cost function of the FCA solution (computed by Equation 4.5), and the number of channel separation violations occurring in the resulting channel assignment.

a. 1st floor



b. 2nd floor



c. 3rd floor

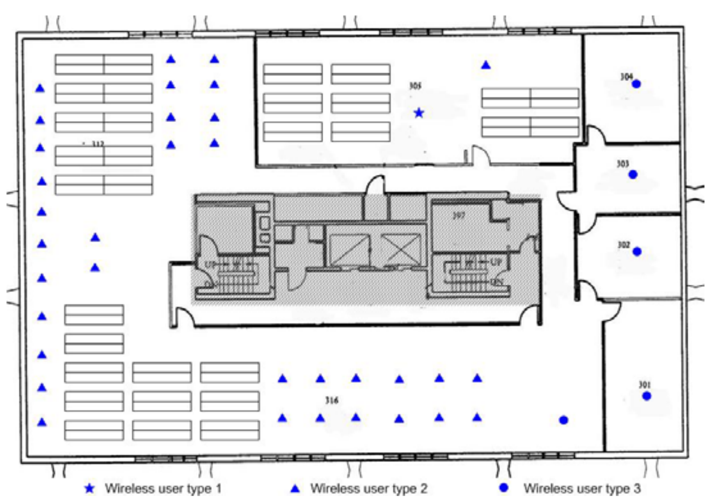
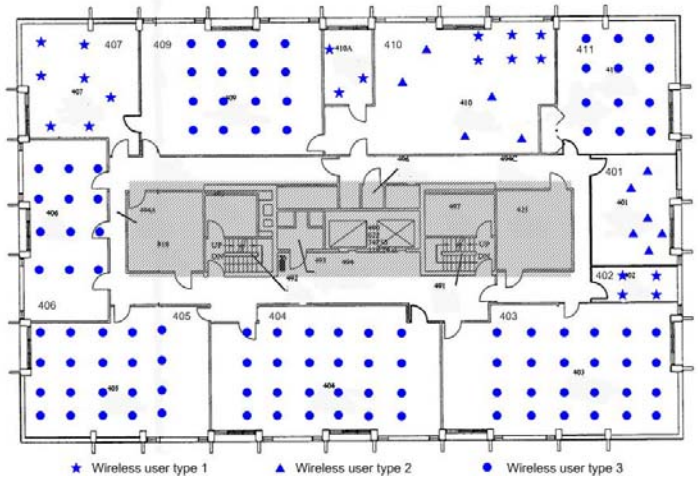


Figure 5.1 Floor plans of the School of Information Science building

d. 4th floor

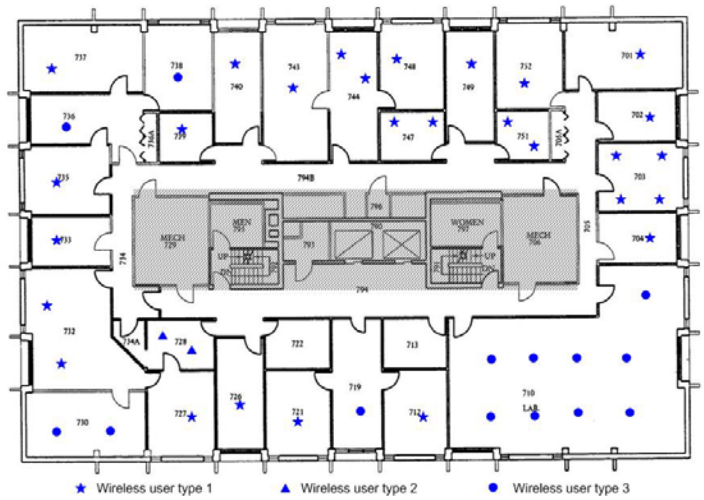


e. 5th floor

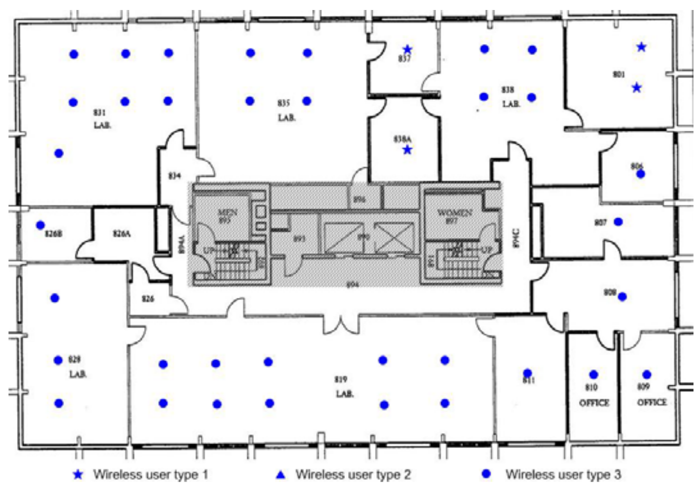
f. 6th floor

Figure 5.1 Floor plans of the School of Information Science building (continued)

f. 7th floor



g. 8th floor



Note: A symbol "Wireless user type *i*" indicates wireless user situating in the sub-areas of type *i* as defined in Table 5.1.

Figure 5.1 Floor plans of the School of Information Science building (continued)

Table 5.1 Network usage characteristics

Sub-areas	User activity level	Average user data rate (Kbps)
Type 1: Private sub-areas (e.g., staff and graduate student offices)	$\alpha_1 = 0.70$	$R_1 = 460$
Type 2: Public sub-areas for unscheduled activities (e.g., library, student lounges)	$\alpha_2 = 0.55$	$R_2 = 260$
Type 3: Public sub-areas for schedule-based activities (e.g,. classrooms, laboratories)	$\alpha_3 = 0.50$	$R_3 = 80$

5.2.3 Numerical Results and Discussion

Tables 5.2 and 5.3 summarize the numerical results of the frequency channel assignment process for Networks 1 and 2, respectively. Tables 5.2a and 5.3a report the cost and the number of channel separation violations of the optimal FCA solution as determined by the exhaustive search. Additionally, they report the CPU time of the exhaustive enumeration. Tables 5.2b through 5.2e and Tables 5.3b through 5.3e show numerical results derived when different values of α are considered. The performance of the implemented FCA process is measured by the cost of the FCA solution, the number of channel separation violations and the CPU time (in seconds). The results displayed in the tables are highlighted if the cost and the violation value are equal to the optimal result obtained in the exhaustive search. For each combination of parameter settings, the simulated annealing process was run ten times. The average value and standard deviation of the cost function, the number of channel separation violations and the CPU time are computed and reported in Tables 5.2 and 5.3. From the results shown in the tables, the following observations are made:

5.2.3.1 Efficiency of the Simulated Annealing Method in the FCA Phase

The results in Tables 5.2 and 5.3 show that the simulated annealing method implemented for the FCA phase can find the optimal solution for the test networks in a much shorter CPU time than the exhaustive search. In the case of Network 1, the exhaustive search took 3 minutes to enumerate all possible solutions and to determine the optimal FCA solution. By comparison, the simulated annealing method spent as little as 0.02 seconds with a proper parameter setting of $\varphi = 0.7$, $P_{a\_init} = 0.3$, and $Max\_noimprov = 100$. In the case of Network 2, the exhaustive search spent 6 hours and 6 minutes to obtain the optimal FCA solution while the implemented simulated annealing method spent as little as 0.06 seconds when parameters were $\varphi = 0.7$, $P_{a\_init} = 0.3$, and $Max\_noimprov = 150$. In some cases in which the value of $Max\_noimprov$ is set too low while the values of φ and P_{a_init} are set too high (such as when $\varphi = 0.9$, $P_{a\_init} = 0.8$, and $Max\_noimprov < 500$) the simulated annealing method was unable to reach the optimal solution and stopped at local optima. Thus, the appropriate parameter settings must be defined for the simulated annealing method to work effectively.

Table 5.2 Sensitivity analysis of control parameters in FCA phase: Network 1

a. Optimal solution

	Cost	# channel separation violation	CPU time
optimal frequency channel assignment obtained from exhausive search	24	12	3 mins

b. $\varphi = 0.3$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	25.20	13.20	0.019
		stdv	1.55	1.55	0.007
	100	avg	24.00	12.00	0.028
		stdv	0.00	0.00	0.008
	150	avg	24.00	12.00	0.027
		stdv	0.00	0.00	0.009
	200	avg	24.00	12.00	0.032
		stdv	0.00	0.00	0.009
	250	avg	24.00	12.00	0.036
		stdv	0.00	0.00	0.007
0.5	50	avg	26.70	14.70	0.011
		stdv	0.95	0.95	0.003
	100	avg	29.40	17.40	0.017
		stdv	1.90	1.90	0.011
	150	avg	24.00	12.00	0.030
		stdv	0.00	0.00	0.006
	200	avg	24.00	12.00	0.034
		stdv	0.00	0.00	0.011
	250	avg	24.00	12.00	0.038
		stdv	0.00	0.00	0.008
0.8	100	avg	29.00	17.00	0.012
		stdv	0.00	0.00	0.004
	200	avg	28.00	16.00	0.022
		stdv	2.11	2.11	0.009
	300	avg	24.00	12.00	0.050
		stdv	0.00	0.00	0.008
	400	avg	24.00	12.00	0.062
		stdv	0.00	0.00	0.006
	500	avg	24.00	12.00	0.063
		stdv	0.00	0.00	0.005

c. $\varphi = 0.5$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	24.00	12.00	0.017
		stdv	0.00	0.00	0.007
	100	avg	24.00	12.00	0.020
		stdv	0.00	0.00	0.005
	150	avg	24.00	12.00	0.027
		stdv	0.00	0.00	0.008
	200	avg	24.00	12.00	0.031
		stdv	0.00	0.00	0.006
	250	avg	24.00	12.00	0.033
		stdv	0.00	0.00	0.005
0.5	50	avg	31.40	18.70	0.011
		stdv	0.97	0.48	0.003
	100	avg	25.50	13.50	0.020
		stdv	2.42	2.42	0.008
	150	avg	24.00	12.00	0.030
		stdv	0.00	0.00	0.000
	200	avg	24.00	12.00	0.032
		stdv	0.00	0.00	0.004
	250	avg	24.00	12.00	0.035
		stdv	0.00	0.00	0.008
0.8	100	avg	31.40	19.40	0.016
		stdv	2.07	2.07	0.008
	200	avg	27.20	15.20	0.020
		stdv	2.35	2.35	0.015
	300	avg	24.00	12.00	0.049
		stdv	0.00	0.00	0.007
	400	avg	24.00	12.00	0.055
		stdv	0.00	0.00	0.007
	500	avg	24.00	12.00	0.066
		stdv	0.00	0.00	0.008

d. $\varphi = 0.7$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	24.00	12.00	0.015
		stdv	0.00	0.00	0.007
	100	avg	24.00	12.00	0.020
		stdv	0.00	0.00	0.005
	150	avg	24.00	12.00	0.025
		stdv	0.00	0.00	0.005
	200	avg	24.00	12.00	0.032
		stdv	0.00	0.00	0.006
	250	avg	24.00	12.00	0.039
		stdv	0.00	0.00	0.006
0.5	50	avg	30.00	18.00	0.012
		stdv	0.00	0.00	0.004
	100	avg	24.00	12.00	0.021
		stdv	0.00	0.00	0.003
	150	avg	24.50	12.50	0.026
		stdv	1.58	1.58	0.007
	200	avg	24.00	12.00	0.033
		stdv	0.00	0.00	0.005
	250	avg	24.00	12.00	0.040
		stdv	0.00	0.00	0.007
0.8	100	avg	27.60	15.60	0.011
		stdv	0.97	0.97	0.003
	200	avg	28.90	16.90	0.013
		stdv	0.32	0.32	0.005
	300	avg	27.30	15.30	0.021
		stdv	1.25	1.25	0.011
	400	avg	24.00	12.00	0.046
		stdv	0.00	0.00	0.005
	500	avg	24.00	12.00	0.068
		stdv	0.00	0.00	0.009

e. $\varphi = 0.9$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	25.50	13.50	0.017
		stdv	2.42	2.42	0.007
	100	avg	27.60	15.60	0.012
		stdv	0.97	0.97	0.006
	150	avg	24.00	12.00	0.026
		stdv	0.00	0.00	0.005
	200	avg	24.00	12.00	0.028
		stdv	0.00	0.00	0.008
	250	avg	24.00	12.00	0.031
		stdv	0.00	0.00	0.003
0.5	50	avg	28.60	16.60	0.012
		stdv	2.07	2.07	0.004
	100	avg	29.00	17.00	0.015
		stdv	0.00	0.00	0.007
	150	avg	24.00	12.00	0.025
		stdv	0.00	0.00	0.005
	200	avg	24.00	12.00	0.021
		stdv	0.00	0.00	0.012
	250	avg	24.00	12.00	0.036
		stdv	0.00	0.00	0.005
0.8	100	avg	29.40	17.20	0.013
		stdv	0.84	0.42	0.005
	200	avg	29.00	17.00	0.013
		stdv	0.00	0.00	0.007
	300	avg	29.00	17.00	0.014
		stdv	0.00	0.00	0.005
	400	avg	29.00	17.00	0.012
		stdv	0.00	0.00	0.004
	500	avg	24.00	12.00	0.058
		stdv	0.00	0.00	0.006

Table 5.3 Sensitivity analysis of control parameters in FCA phase: Network 2

a. Optimal solution

	Cost	# channel separation violation	CPU time
optimal frequency channel assignment obtained from exhausive search	102	51	6 hours. 6 mins

b. $\varphi = 0.3$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	113.00	62.00	0.024
		stdv	0.00	0.00	0.007
	100	avg	114.60	63.60	0.046
		stdv	10.84	10.84	0.020
	150	avg	102.00	51.00	0.094
		stdv	0.00	0.00	0.011
	200	avg	102.00	51.00	0.121
		stdv	0.00	0.00	0.014
	250	avg	102.00	51.00	0.143
		stdv	0.00	0.00	0.012
0.5	50	avg	113.60	62.60	0.020
		stdv	0.52	0.52	0.009
	100	avg	113.00	62.00	0.031
		stdv	0.00	0.00	0.007
	150	avg	109.70	58.70	0.046
		stdv	5.31	5.31	0.026
	200	avg	109.20	58.20	0.058
		stdv	6.20	6.20	0.039
	250	avg	102.00	51.00	0.145
		stdv	0.00	0.00	0.011
0.8	300	avg	118.50	67.50	0.038
		stdv	5.80	5.80	0.008
	400	avg	109.20	58.20	0.097
		stdv	6.20	6.20	0.065
	500	avg	105.30	54.30	0.200
		stdv	5.31	5.31	0.098
	600	avg	102.00	51.00	0.282
		stdv	0.00	0.00	0.032
	700	avg	102.00	51.00	0.324
		stdv	0.00	0.00	0.034

c. $\varphi = 0.5$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	111.90	60.90	0.022
		stdv	3.48	3.48	0.010
	100	avg	102.00	51.00	0.068
		stdv	0.00	0.00	0.014
	150	avg	102.00	51.00	0.093
		stdv	0.00	0.00	0.009
	200	avg	102.00	51.00	0.101
		stdv	0.00	0.00	0.006
	250	avg	102.00	51.00	0.122
		stdv	0.00	0.00	0.006
0.5	50	avg	116.00	65.00	0.021
		stdv	4.83	4.83	0.006
	100	avg	110.50	59.50	0.026
		stdv	2.64	2.64	0.011
	150	avg	107.40	56.40	0.071
		stdv	9.03	9.03	0.031
	200	avg	106.40	55.40	0.092
		stdv	9.28	9.28	0.036
	250	avg	102.00	51.00	0.129
		stdv	0.00	0.00	0.013
0.8	300	avg	113.60	62.60	0.040
		stdv	1.90	1.90	0.005
	400	avg	116.00	65.00	0.044
		stdv	3.16	3.16	0.013
	500	avg	106.80	55.80	0.166
		stdv	6.20	6.20	0.096
	600	avg	102.00	51.00	0.273
		stdv	0.00	0.00	0.019
	700	avg	102.00	51.00	0.329
		stdv	0.00	0.00	0.017

d. $\varphi = 0.7$

P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	111.30	60.30	0.020
		stdv	7.82	7.82	0.007
	100	avg	106.30	55.40	0.035
		stdv	7.44	7.55	0.014
	150	avg	102.00	51.00	0.064
		stdv	0.00	0.00	0.008
	200	avg	102.00	51.00	0.072
		stdv	0.00	0.00	0.008
	250	avg	102.00	51.00	0.086
		stdv	0.00	0.00	0.013
0.5	50	avg	118.10	66.40	0.014
		stdv	4.56	4.77	0.005
	100	avg	113.90	65.60	0.021
		stdv	3.67	5.74	0.006
	150	avg	112.30	61.10	0.028
		stdv	6.13	5.86	0.021
	200	avg	108.70	57.70	0.052
		stdv	9.03	9.03	0.027
	250	avg	102.00	51.00	0.087
		stdv	0.00	0.00	0.012
0.8	300	avg	115.10	64.10	0.028
		stdv	4.43	4.43	0.009
	400	avg	116.80	65.90	0.030
		stdv	4.87	5.04	0.007
	500	avg	108.60	57.60	0.110
		stdv	9.16	9.16	0.069
	600	avg	107.50	56.50	0.132
		stdv	7.78	7.78	0.084
	700	avg	102.00	51.00	0.218
		stdv	0.00	0.00	0.014

e. $\varphi = 0.9$

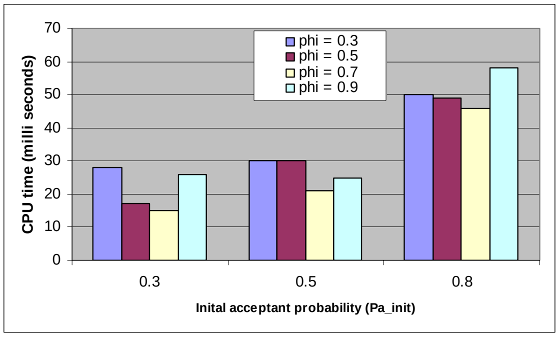
P_{a_init}	*Max_noimprov*	statistic	Cost	# channel separation violation	CPU time (sec)
0.3	50	avg	114.60	63.50	0.015
		stdv	6.26	6.19	0.007
	100	avg	113.10	62.00	0.020
		stdv	6.44	6.27	0.005
	150	avg	113.10	62.10	0.032
		stdv	7.03	7.03	0.018
	200	avg	103.70	52.60	0.071
		stdv	5.38	5.06	0.017
	250	avg	102.00	51.00	0.096
		stdv	0.00	0.00	0.010
0.5	100	avg	117.10	66.10	0.023
		stdv	5.30	5.30	0.008
	200	avg	115.00	64.00	0.019
		stdv	4.22	4.22	0.003
	300	avg	110.80	59.80	0.040
		stdv	4.64	4.64	0.035
	400	avg	107.50	56.60	0.092
		stdv	7.78	8.02	0.052
	500	avg	102.00	51.00	0.167
		stdv	0.00	0.00	0.020
0.8	600	avg	118.70	67.50	0.034
		stdv	6.29	6.60	0.005
	800	avg	117.10	66.10	0.044
		stdv	5.30	5.30	0.005
	1000	avg	113.60	62.60	0.055
		stdv	1.90	1.90	0.012
	1200	avg	109.20	58.30	0.209
		stdv	8.15	8.34	0.155
	1400	avg	102.00	51.00	0.409
		stdv	0.00	0.00	0.036

5.2.3.2 Effects of P_{a_init} on the FCA Phase

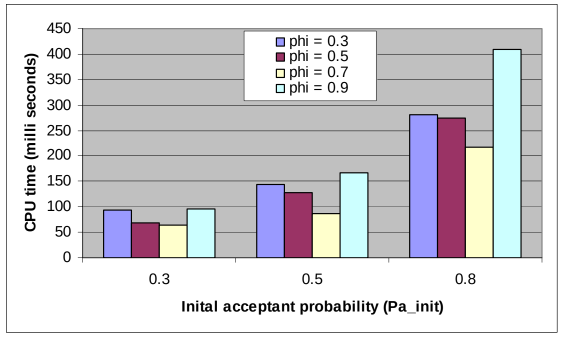
For each value of φ considered in the study, as the value of P_{a_init} increases, the simulated annealing method searches an increasing number of iterations before converging to the optimal solution. In other word, a high value of $Max\_noimprov$ needs to be set in order to allow the search to reach the optimal solution. This results in a longer CPU time for the simulated annealing method to converge and terminate. Figure 5.2 summarizes the effects of P_{a_init} on the CPU time. The figure plots the CPU time versus P_{a_init} with different values of φ for Networks 1 and 2. Consider the results shown for Network 1 in Figure 5.2a. We can see that P_{a_init} values of 0.3 and 0.5 result in shorter CPU times compared to those when P_{a_init} is 0.8 for all values of φ considered. As the problem size increases, it is obvious that the small values of P_{a_init} outperform the larger value, as we can observe in the numerical results of the Network 2 study (Figure 5.2b). The P_{a_init} of 0.3 yields a more efficient simulated annealing method in terms of CPU time than all values of φ considered.

The reason for these phenomena is that when the values of P_{a_init} are high, the search begins with a high probability to accept inferior neighbor solutions. In the implemented frequency channel assignment process, a good initial FCA solution is derived and input to the search process, indicating that the search began at a good quality solution. As a result, setting a high value for the initial acceptance probability causes the search to waste time accepting many unnecessary, poor solutions and to spend more time before converging on a solution. In the case when a good initial solution is provided, the starting value of the acceptance probability (which, in turn determines an appropriate value of CT_{max} as computed by Equation 5.2) should be set at a low value.

From these observations, we conclude that a low value of P_{a_init} is suitable for the frequency channel assignment process when a good initial assignment is provided for the simulated annealing method. Thus, P_{a_init} of 0.3 is used in further numerical experiments.



a. Network 1



b. Network 2

Figure 5.2 Effect of P_{a_init} on the computational time of the simulated annealing method

5.2.3.3 Effects of φ on the FCA Phase

For the simulated annealing method to be efficient, φ needs to be set at a proper value. Too small or too high a value of φ could result in a longer search time during the frequency channel assignment process. In the following sensitivity analysis study, φ values of 0.3, 0.5, 0.7, and 0.9 are tested. As shown in the numerical results presented in Tables 5.2 and 5.3 and summarized in Figure 5.2, φ of 0.7 yields the most efficient Simulate Annealing method, converging to the optimal solution in less CPU time for all values of P_{a_init} in both test networks.

The reasons for these results and observations are as follows: φ determines a decreasing rate of the control temperature and an increasing rate of the number of iterations at each temperature step. In turn, the control temperature defines the probability of accepting an inferior FCA solution. If the control temperature decreases too fast (with a small value of φ), the search might miss chances to explore some inferior solutions which might lead to the optimal solution because the probability of accepting inferior solutions also decreases as the control temperature decreases. If the control temperature decreases too slow (with a high value of φ), the search might accept too many unnecessary inferior solutions resulting in slow convergence of the search.

Consider the number of iterations at each temperature step. Typically, as the search progresses and as the control temperature decreases, the number of iterations at each temperature step increases to allow the search to explore the neighborhood more thoroughly as it approaches the optimal solution. However, if the number of iterations at each temperature step increases at a high rate (with a small value of φ), the search can remain at a particular temperature for too long. If the number of iterations at each temperature step increases at a low rate (with a high value of φ), the search might move to the next temperature step too fast. Thus, the appropriate value of φ must be set to achieve a balance between the decreasing rate of the control temperature and the increasing rate of the number of iterations at each temperature step.

From the numerical results and observations, φ value of 0.7 results in the most efficient simulated annealing method and this value will be used in further numerical experiments.

5.2.3.4 Effects of stopping criteria on the FCA phase

The numerical results show that in all cases the simulated annealing method terminated the search because it reached the *Max_noimprov* stopping criteria. The search stopped before the control temperature reaches the CT_{min} of 0.01. This means the CT_{min} of

0.01 is low enough to allow the simulated annealing search process to reach a point of equilibrium.

The value of $Max\_noimprov$ should be set such that the search has a sufficient number of iterations to explore the solution space and could converge to the optimal (or near optimal) solution without wasting too much time after reaching the optimal solution before terminating the search process. From the numerical results in Tables 5.2 and 5.3, with values of $\varphi = 0.7$ and $P_{a\_init} = 0.3$. $Max\_noimprov$ of 150 is sufficient to allow the search to converge to the optimal solution in both test networks. Thus, the value of $Max\_noimprov$ is set at 150 in further numerical experiments.

5.2.4 Parameter Setting of the Simulated Annealing Method in the FCA Phase

In summary, parameter settings of the simulated annealing method implemented for the FCA phase that will be used in further experiments are: $\varphi = 0.7$, $P_{a\_init} = 0.3$, $CT_{min} = 0.01$ and $Max\_noimprov = 150$.

5.3 Sensitivity Analysis on Static Parameters in the CVR Phase

The Constraint Violation Reduction (CVR) phase aims to adjust the APs' locations and power levels in order to improve network performance and obtain a network configuration that meets the design criteria. The search procedures of the CVR phase use constraint violation information to guide the search and utilize the memory techniques based on the tabu search method to prevent the occurrence of cycling in the search.

The constraint violation information is evaluated and recorded for each predefined area, called a unit area (UA). Such information provides useful data that aids in the indication of potential parameters' adjustments that may lead to improvement of network performance. The memory techniques are implemented such that the most recently moved AP is not allowed to move back in the direction from which it originated for an allotted period of time (i.e., a number of iterations called the tabu tenure). The memory approach is also applied in changing the AP's power level parameter. The CVR process continues until either a feasible solution is obtained or a stopping condition is reached.

The main parameters involved in the CVR process are the size of the unit area (UA) for managing and recording the constraint violation information, the tabu tenure (tabu list length) for both the AP's move direction and power level attribute, and the number of consecutive non-improving iterations (Max_{iter_CVR}) which will cause the CVR process to stop. In the following sections, we present the results of a sensitivity analysis of these parameters and determine the appropriate values that yield an efficient CVR process.

5.3.1 Parameter Settings for the Sensitivity Analysis

We conducted a sensitivity analysis on the key parameters of the CVR process, including the length of tabu tenures, the Max_{iter_CVR} and the size of the unit area (UA). We test the following parameter settings:

- The length of tabu tenure can be either fixed (static tabu tenure) or varied (dynamic tabu tenure) during the course of the search. In this study, we examine both static and dynamic tabu tenures for an AP's move direction and power level attribute. Several combinations of static tabu tenure for both attributes are tested. For both the move direction attribute and the power level attribute, tabu list sizes of 1, 4, and 8 are tested. In addition, we examine dynamic tabu tenures in which the tabu list size of the move direction attribute is randomly selected from the ranges [2, 4] and [4, 8], while the tabu list size of the power level attribute are varied from the ranges [1, 4] and [4, 8]. These values are selected among several values considered in the preliminary tests. The selected values indicate good potential for the tabu search and are thoroughly explored in the parameter tuning test presented in the following sections.
- For parameter Max_{iter_CVR}, we test 30, 50, 100, and 150.
- The UA size is set relative to the size of coverage square (CS) which is used to estimate a coverage area of an AP as described in section 4.2.1. The UA is set at half of the CS size (0.5×CS), equal to the CS size (1×CS) and twice the CS size (2×CS). The effects of the different UA sizes on efficiency of the CVR process are investigated.

5.3.2 Test Scenarios

In order to study the effects of the parameters described above, we conducted extensive numerical experiments of WLAN design on several scenarios. We considered network service environments at the School of Information Science (SIS) building and the Hillman library (HL) building at the University of Pittsburgh. Four design scenarios, ranging from the simple case of a small, single-floor network design to the complex case of a service area covering multiple floors are considered.

Scenario 1 considers a single-floor design in which the target service area is located on the fourth floor of the SIS building (called SIS4). Scenario 2 considers a two-floor design in which the target service area covers both the fourth and fifth floors of the SIS building. Scenario 3 considers a four-floor design in which the target service area ranges from the first floor through the fourth floor of the SIS building. Finally, Scenario 4 is a single-floor design considering the particular WLAN service needs for the first floor of the HL building (called HL1). The floor plan of the SIS building is shown in Figure 5.1. Figure 5.3 shows the floor plan of the first floor of Hillman Library. It is assumed that traffic activities of wireless users follow the description in Table 5.4.

Table 5.4 Network usage characteristics

Sub-areas	User activity level	Average user data rate (Kbps)
Type 1: Private sub-areas (e.g., library staff offices, graduate student offices)	$\alpha_1 = 0.50$	$R_1 = 460$
Type 2: Public sub-areas for unscheduled activities (e.g., library study areas, student lounges)	$\alpha_2 = 0.40$	$R_2 = 260$
Type 3: Public sub-areas for schedule-based activities (e.g., classrooms, laboratories, meeting rooms)	$\alpha_3 = 0.35$	$R_3 = 80$

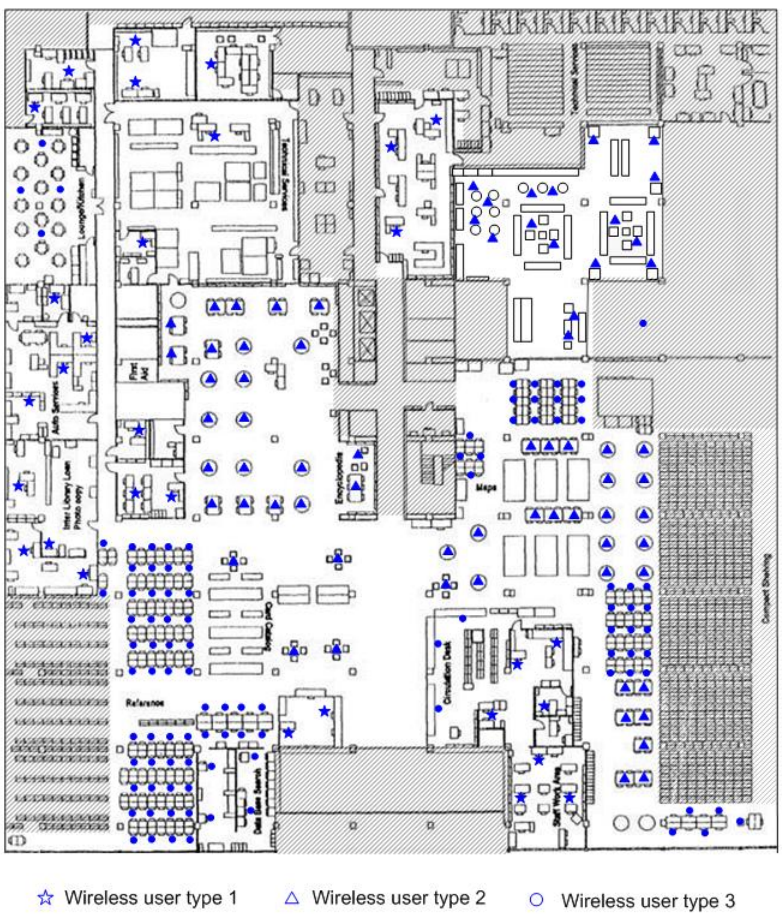


Figure 5.3 Floor-plan and demand node distribution for the first floor of Hillman Library (HL1)

The sensitivity analysis study described in this section focuses on the efficiency of the CVR process as developed based on the tabu search approach with a special implementation of constraint violation information to help guild the search. The performance measures of this sensitivity analysis include CPU time and the constraint violation value of the obtained WLAN configuration (see section 4.4.3 for the evaluation function). The design process starts by generating an initial configuration in which the number of APs and their locations and power levels are derived from the construction phase (see section 4.2) and their frequency channels are derived from the FCA phase (see section 4.3). The CVR process is run until either a feasible configuration is obtained or until a stopping rule is reached, whichever come first. In the latter case, an intensification phase is activated in which the CVR process is restarted and reinitialized with the first solution on the elite list. This procedure of restarting the CVR process is repeated until a feasible solution is obtained or until the elite list is depleted.

To test the CVR process thoroughly, for each design scenario and combination of parameter settings, the solution technique is run ten times. Average values and standard deviations of the CPU time and the constraint violation values of the resulting WLAN configurations are reported.

5.3.3 Numerical Results and Discussion

5.3.3.1 Effects of Tabu Tenure Length on the CVR Phase

Table 5.5 compares the results obtained by using the static tabu tenure with the results obtained by using the dynamic tabu tenure in various network design scenarios. The table reports the average and the standard deviation of the resulting constraint violation values as well as the CPU times. Findings show that use of the dynamic tabu tenure outperforms the static tabu tenure in the search process. In all scenarios tested, the search using dynamic tabu tenures converges to a feasible solution faster than a search using static tabu tenures. This is because the dynamic tabu tenure more effectively prevents cycling in the search. When the static tabu tenure is used, small tabu tenure periodically can cause a repeat of the solutions in the search. Tabu tenures that are too large can prevent the search from exploring some regions that may represent vicinities of good solutions. The dynamic tabu tenure avoids looping in the search and directs it to different parts of the solution space.

As shown in Table 5.5, all ranges of tabu list size for the dynamic tabu tenure perform relatively well. Especially with the range [4, 8] for the direction attribute and the range [1, 4] for the power level attribute, the search yields good results, converging to a feasible solution relatively quickly, for all design scenarios considered. Therefore, we will use a dynamic tabu tenure with the range [4, 8] for the direction attribute and the range [1, 4] for the power level attribute for the remainder of the numerical experiments.

Table 5.5 Effects of tabu tenure on the CVR process

Design scenario	Tabu tenure			Constraint violation		CPU time (sec)	
	Type	AP's move direction attribute	power level attribute	avg	stdv	avg	stdv
scenario 1 (SIS4)	static	1	1	0.000	0.000	0.327	0.005
		4	4	0.000	0.000	0.325	0.002
		8	8	0.000	0.000	0.325	0.003
	dynamic	(2,4)	(1,4)	0.000	0.000	0.322	0.001
			(4,8)	0.000	0.000	0.323	0.004
		(4,8)	(1,4)	0.000	0.000	0.326	0.008
			(4,8)	0.000	0.000	0.322	0.002
scenario 2 (2 floors)	static	1	1	0.000	0.000	8.20	4.20
		4	4	0.000	0.000	9.52	3.15
		8	8	0.000	0.000	14.66	7.04
	dynamic	(2,4)	(1,4)	0.000	0.000	8.03	2.8
			(4,8)	0.000	0.000	8.42	3.4
		(4,8)	(1,4)	0.000	0.000	7.34	3.5
			(4,8)	0.000	0.000	7.45	2.0
scenario 3 (4 floors)	static	1	1	0.000	0.000	118.57	38.62
		4	4	0.000	0.000	149.45	78.25
		8	8	0.000	0.000	170.71	96.31
	dynamic	(2,4)	(1,4)	0.000	0.000	108.80	47.5
			(4,8)	0.000	0.000	120.91	42.5
		(4,8)	(1,4)	0.000	0.000	93.86	18.0
			(4,8)	0.000	0.000	147.79	36.6
scenario 4 (HL1)	static	1	1	0.000	0.000	99.57	46.1
		4	4	0.000	0.000	115.01	42.9
		8	8	0.000	0.000	127.43	42.1
	dynamic	(2,4)	(1,4)	0.000	0.000	96.88	36.8
			(4,8)	0.000	0.000	100.93	75.9
		(4,8)	(1,4)	0.000	0.000	83.64	34.9
			(4,8)	0.000	0.000	92.46	65.5

5.3.3.2 Effect of Max_{iter_CVR} on the CVR Phase

To study the effects of the number of consecutive iterations that the search can proceed without improvement, the dynamic tabu tenure and the UA size of 1xCS are used. The Max_{iter_CVR} is set at 30, 50, 100, and 150 iterations. The tests are run for network design scenario 1 through 4. The results are summarized in Table 5.6.

Table 5.6 Effects of Max_{iter_CVR} on the CVR process

Design scenario	# iterations without improvement (Max_{iter_CVR})	Constraint violation	
		avg	stdv
scenario 1 (SIS4)	30	0.0000000	0.0000000
	50	0.0000000	0.0000000
	100	0.0000000	0.0000000
	150	0.0000000	0.0000000
scenario 2 (2 floors)	30	0.0000061	0.0000183
	50	0.0000000	0.0000000
	100	0.0000000	0.0000000
	150	0.0000000	0.0000000
scenario 3 (4 floors)	30	0.0002051	0.0005975
	50	0.0001192	0.0003409
	100	0.0000000	0.0000000
	150	0.0000000	0.0000000
scenario 4 (HL1)	30	0.0000330	0.0000536
	50	0.0000257	0.0000501
	100	0.0000000	0.0000000
	150	0.0000000	0.0000000

Table 5.6 reports the averages and standard deviations of the constraint violations for each of the four design scenarios. We can see that with a small Max_{iter_CVR} of 30, the search cannot obtain a feasible solution in the multi-floor design scenarios (scenarios 2 and 3) or the large service area (scenario 4). With Max_{iter_CVR} of 50, the search cannot obtain a feasible solution for scenarios 3 and 4. However, the search can find feasible solutions for all design scenarios when the value of Max_{iter_CVR} is set at 100 or 150. The Max_{iter_CVR} of 100 is sufficient to allow the search to converge to a feasible solution within a reasonable CPU time. Therefore, a Max_{iter_CVR} of 100 will be used in further numerical experiments.

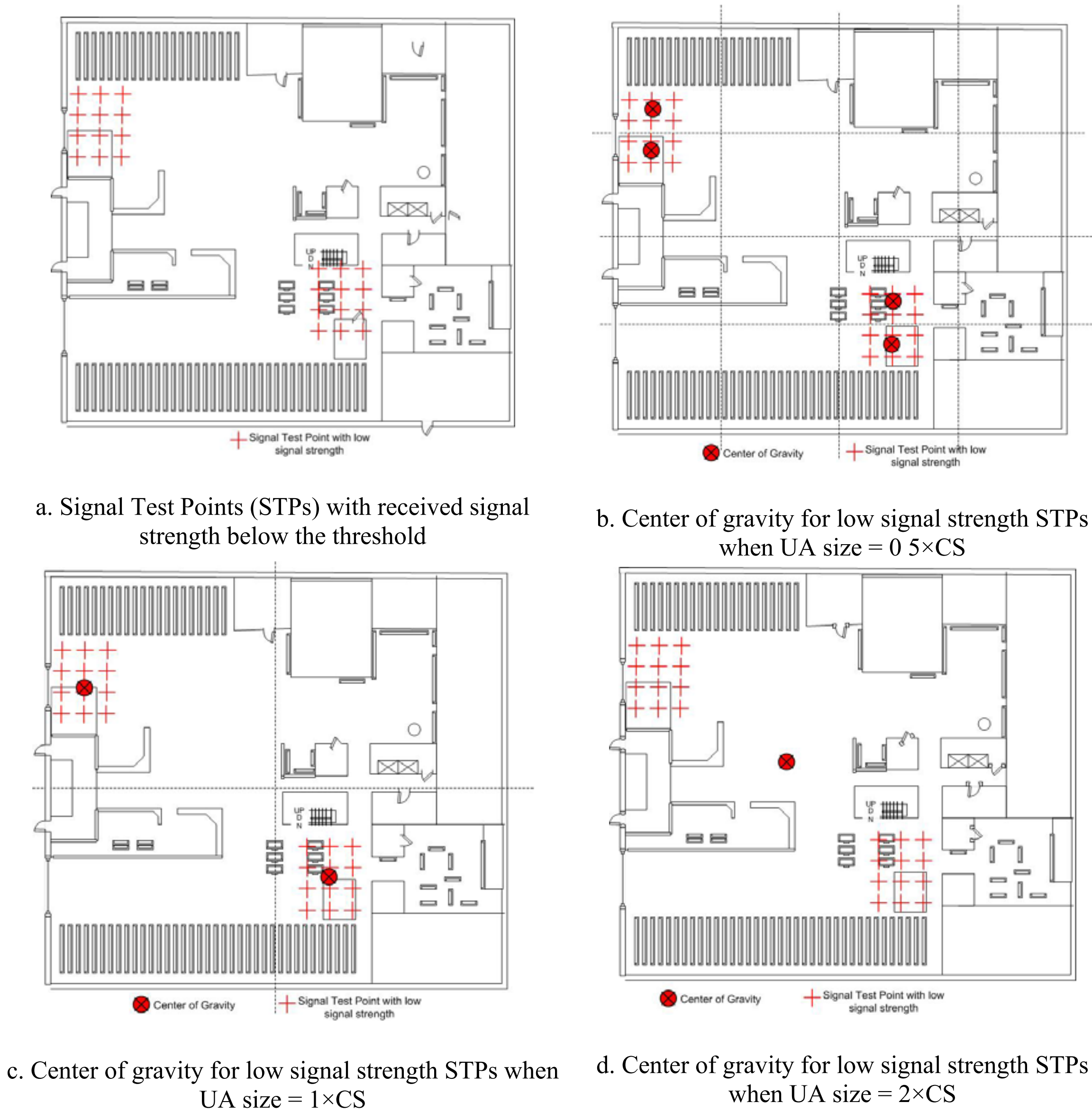
5.3.3.3 Effect of the UA Size on the CVR Phase

In order to manage the constraint violation information that guides the search in the CVR process, the service area is divided into equally sized areas called unit areas (UA) in which the received signal strength and the SIR level of the signal test points (STPs) are evaluated. In case any performance criteria fail to meet the design requirement, the constraint violation are calculated and recorded in each UA. In this section, the effects of UA sizes on the efficiency of the CVR process are studied.

Table 5.7 reports the averages and standard deviations of CPU times and the constraint violations of the solutions obtained as a result of the search process. As shown in the table, the search using a UA size of 1×CS (i.e., the size of the coverage square) converges to a feasible solution faster than any of the other sizes. When the UA size is smaller or bigger than 1×CS (i.e., 0.5×CS or 2×CS), the search takes a longer time to reach a feasible solution and, in some cases, is not able to reach a feasible solution. The reason is that too small or too big a unit area may provide incorrect constraint violation information. Figure 5.4 illustrates such cases. Figure 5.4a shows the STPs in the service area that have signal strength below the $P_{Rthreshold}$. Figures 5.4b, 5.4c and 5.4d show Center of Gravities (CoGs) of such STPs with UA sizes of 0.5×CS, 1×CS, and 2×CS, respectively. When a small UA size is used, the service area is divided into several sections to manage the constraint violation information. Constraint information of the UA

Table 5.7 Effects of UA size on the efficiency of the CVR process

Design scenario	Size of unit area	CPU time (sec)		Constraint violation	
		avg	stdv	avg	stdv
scenario 1 (SIS4)	0.5×CS	8.156	1.295	0.0108	0.0040
	1×CS	0.306	0.010	0.0000	0.0000
	2×CS	0.248	0.003	0.0000	0.0000
scenario 2 (2 floors)	0.5×CS	25.715	11.470	0.0327	0.0091
	1×CS	7.214	3.686	0.0000	0.0000
	2×CS	138.056	38.067	0.0072	0.0057
scenario 3 (4 floors)	0.5×CS	527.080	130.429	0.7466	0.0173
	1×CS	93.203	24.267	0.0000	0.0000
	2×CS	1024.295	373.893	0.0064	0.0039
scenario 4 (HL1)	0.5×CS	357.054	194.320	0.0022	0.0023
	1×CS	84.646	34.974	0.0000	0.0000
	2×CS	128.723	95.263	0.0006	0.0011



a. Signal Test Points (STPs) with received signal strength below the threshold

b. Center of gravity for low signal strength STPs when UA size = 0 5×CS

c. Center of gravity for low signal strength STPs when UA size = 1×CS

d. Center of gravity for low signal strength STPs when UA size = 2×CS

Figure 5.4 Examples of problems in small or large UA size

is then used in the move operation to determine APs and their parameters that should be adjusted to improve the solution. If one uses a small UA size, too many unnecessary CoGs may result as shown in Figure 5.4b. In this case, a group of problematic STPs may be represented with too many CoGs. When the UA is too large (Figure 5.4d), the resulting CoG may provide incorrect information about the constraint violation and mislead the search. This results in a longer search time before a feasible solution is found. In some cases, the search may never be able to reach a feasible solution.

5.3.3.4 Effect of Step size (λ) on the CVR Phase

APs' locations are adjusted as they are move in a direction specified by the move operator for λ step size at a time. A small step size may result in it taking a long time for an AP to reach a desired location while a large step size may be too coarse and move the AP further away from the desired location. As a result, it might take a longer time for the AP to reach a feasible location and the search to end. Table 5.8 reports the average and standard deviation of the CPU time and the constraint violation values. When step sizes of 3 and 5 meters are used, there is a longer search time and, in some cases, the search is unable to obtain a feasible solution. A step size of 1 meter works well in all design scenarios tested. Thus, a step size of 1 meter is used in further experiments.

Table 5.8 Effects of step size (λ) on the CVR process

Design scenario	Step size	CPU time (sec)		Constraint violation	
		avg	stdv	avg	stdv
scenario 1 (SIS4)	1 m.	0.306	0.010	0.000000	0.000000
	3 m.	0.673	0.335	0.000000	0.000000
	5 m.	0.790	0.759	0.000000	0.000000
scenario 2 (2 floors)	1 m.	7.214	3.686	0.000000	0.000000
	3 m.	48.587	55.143	0.000016	0.000032
	5 m.	124.447	106.023	0.000807	0.001285
scenario 3 (4 floors)	1 m.	93.203	24.267	0.000000	0.000000
	3 m.	321.078	191.652	0.001140	0.001825
	5 m.	439.977	164.295	0.002811	0.002913
scenario 4 (HL1)	1 m.	84.646	34.974	0.000000	0.000000
	3 m.	139.317	120.506	0.000000	0.000000
	5 m.	221.468	147.322	0.000005	0.000014

5.3.4 Parameter Setting for the CVR Phase

In summary, parameter settings for the CVR phase that will be used in the rest of numerical experiments are as follows:

- Tabu list size:
    - AP's move direction attribute: dynamic tabu tenure, uniformly selected from 4 to 8
    - Power level attribute: dynamic tabu tenure, uniformly selected from 1 to 4
- Max_{iter_CVR} = 100
- UA size = 1×CS
- Step size to move AP = 1 meter

5.4 Effectiveness of the Heuristic Solution Technique

The effectiveness of the heuristic solution technique in solving the demand-based WLAN design problem is evaluated in two ways: an exhaustive enumeration test and a comparison with other solution techniques.

5.4.1 Exhaustive Enumeration Test

A goal of demand-based WLAN design is to find a network configuration that can satisfy all network design requirements for a particular design scenario. One may ask how difficult it is to find a single feasible solution for a network design problem which is formulated as a Constraint Satisfaction Problem (CSP). If there are a lot of feasible solutions in the search space, such problems may not be difficult and a simple solution technique might be able to find a feasible solution easily.

This section aims to evaluate the difficulty of the demand-based WLAN design problem. The problem difficulty is measured by the tightness of the CSP, defined as the ratio of the number of feasible solutions to the number of all possible solutions in the search space [52].

The exhaustive enumeration was performed over all possible combinations of APs' parameters in the WLAN design of the fourth floor of the School of Information Science (SIS) building for scenario 1 as described in section 5.3.2. The parameters used in the network design are described in Table 6.2. In this design scenario, three APs are required. To determine a network configuration, each AP is pre-assigned a different frequency channel. (Note that channels 1, 6, and 11 are used). This results in a reduced form of the problem in which we need to determine locations and power levels for three APs. There are six power levels that can be assigned to each AP, and the APs can be placed at a grid spacing of 3m × 3m instead of 1m × 1m due to the computationally intensive nature of the exhaustive enumeration. The APs are not allowed to be placed within a specified non-usage area (the central of the floor, housing elevator shafts, stairways and restrooms).

In this problem, there are 98,611,128 possible combinations in which APs' parameters (location and power level) can legally be assigned. At 0.01 seconds per solution, the Sun Blade 1000 workstation with a 750 MHz processor and 2 gigabytes of memory took 273 hours 55 minutes to perform calculations and evaluate the solution space. This number grows exponentially as the problem size enlarges due to either a larger service area and/or higher traffic demand that results in more APs used in the network. Table 5.9 summarizes the results of the exhaustive enumeration. 12,947 solutions are identified as feasible configurations, satisfying all design requirements.

Table 5.9 Summary of the exhaustive enumeration test

Total possible solutions in the search space	98,611,128
Total number of feasible solutions	12,947
Tightness of the problem	0.013%
Total number of feasible solutions when the receiver sensitivity threshold is reduced by 1 dB	24,589
Tightness of the problem when the receiver sensitivity threshold is reduced by 1 dB	0.025%

The results of the exhaustive enumeration indicate that only 0.013% of all possible solutions satisfy all of the design requirements thus making them feasible solutions. There are only 0.025% of all possible solutions that are feasible when the receiver sensitivity threshold is reduced by 1 dB. Therefore, the search space for this particular network design problem is very tight, indicating that the demand-base WLAN design problem is a difficult problem to solve.

5.4.2 Comparison with other Solution Techniques

The demand-based WLAN design problem is a difficult problem for which the computational complexity grows exponentially as the problem size enlarges. According to Prommak et al. [68], the brute-force search technique was preliminary implemented to solve the demand-based WLAN design problem. Such technique is based on enumerating all possible solutions in the search space to find the first feasible solution. Here, the brute-force search changes value of each variable one at a time. In this dissertation, an efficient heuristic solution technique was developed. It utilizes the constraint violation information to help guide the search to those areas where feasible solutions are likely to appear. To evaluate the efficiency of the developed heuristic solution technique, this section compares the proposed solution technique with the brute-force search approach over the four network design scenarios described in section 5.3.2.

Table 5.10 shows the computational results and compares the solution techniques. It details how many times the network evaluation function (Equation 4.9) was executed as well as the overall CPU time spent during the search process. From Table 5.10 we can see that the proposed heuristic solution technique can efficiently solve all the network design problems considered in a much shorter time than the brute-force search technique can. For the SIS4 design scenario (scenario 1), the average CPU time and standard deviation that the proposed heuristic solution technique spent in the search process until finding a feasible network configuration are 0.33 and 0.1 seconds, respectively, while the corresponding average number of function evaluations and standard deviations are 50.57 and 2.74, respectively. By comparison, the brute-force search technique spent 3.5 seconds executing the network evaluation function 546 times. The effectiveness and superiority of the proposed heuristic solution technique becomes more pronounced as larger and more complex network design problems are considered, as demonstrated in the 2-floor, 4-floor and HL1 design scenario (scenarios 2 through 4). In the 2-floor design scenario, the average CPU time and standard deviation of the proposed heuristic solution technique are 7.73 and 4.13 seconds, with a corresponding average number of function evaluations and standard deviation of 299 and 162.44, respectively. By comparison, the brute-force search approach took 157 seconds with the number of function evaluations of 4,899. In the case of the 4-floor and HL1 design scenarios, the brute-force search technique was unable to find a feasible network configuration after evaluating 500,000 and 100,000 network configurations with CPU times of 63,038 and 10,239 seconds, respectively. By comparison, the proposed heuristic solution technique identified feasible network configurations for the 4-floor and HL1 design scenarios after spending an average of 91.09 and 84.06 seconds, respectively, with an average number of function evaluations of 757.93 (4-floor) and 863.36 (HL1).

5.5 Summary

This chapter presents sensitivity analysis and an effectiveness evaluation of the heuristic solution technique developed to solve the demand-based WLAN design problem. Findings from the sensitivity analysis provide appropriate parameter settings to be used in the heuristic solution technique. Such parameters are applied in the numerical experiments conducted in the next chapter. The effectiveness evaluations were conducted in two ways: exhaustive enumeration and comparison with brute-force search techniques. Exhaustive enumeration shows that the demand-based WLAN design problem formulated as the constraint satisfaction problem has a tight search space which is difficult to solve. Comparison with a brute-force search technique shows that the developed heuristic solution technique can efficiently solve the demand-based WLAN CSP in much shorter CPU times compared to the brute-force search techniques.

Table 5.10 Comparison of solution techniques for the demand-based WLAN design problem

Design scenarios	# APs used in network	Brute-force search			Heuristic solution technique				
		Satisfy all design requirements	# function evaluations	CPU time (seconds)	Satisfy all design requirements	# function evaluations		CPU time (seconds)	
						avg	stdv	avg	stdv
scenario 1 (SIS4)	3	Yes	546	3.5	Yes	50.57	2.74	0.33	0.01
scenario 2 (2 floors)	5	Yes	4899	157	Yes	299.00	162.44	7.73	4.13
scenario 3 (4 floors)	9	No (0.0003)*	500,000	63038	Yes	757.93	122.04	91.09	17.36
scenario 4 (HL1)	7	No (0.0003)*	100,000	10239	Yes	863.36	197.73	84.06	31.30

Note: ***** indicate the constraint violation of the best solution found

6.0 NUMERICAL EXPERIMENTS

Chapter 6 presents numerical studies on several aspects of network planning for WLANs. The first part of the chapter explores various scenarios of WLAN design ranging from a single floor WLAN design (section 6.1) to a multi-floor WLAN design (section 6.2) to the design for an area that includes an outdoor courtyard around a building (section 6.3). The second part of this chapter is a sensitivity analysis of several aspects: Section 6.4 presents a numerical study of the effects of user activity level on network configurations and provides a comparative performance evaluation of the demand-based WLAN design model versus coverage-based design. Section 6.5 presents a numerical study of the effects of fading margins on signal coverage availability and network configurations obtained from the design process. Section 6.6 presents a numerical study of the effects of using different path loss models on network configurations obtained from the design.

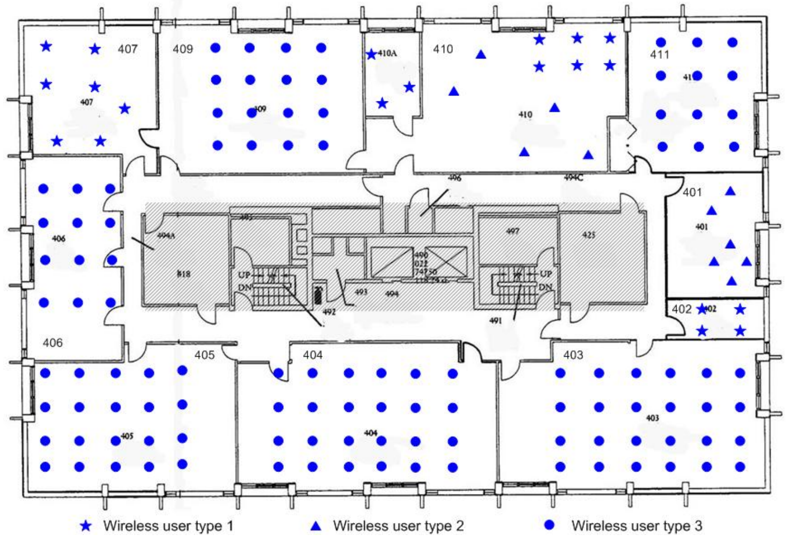
6.1 Single Floor WLAN Designs

In this section, we illustrate WLAN design for a single floor service environment. A variety of service scenarios, including both small and large service areas, are considered. Symbols and definitions of figures used in the presentation of the resulting network configurations are explained.

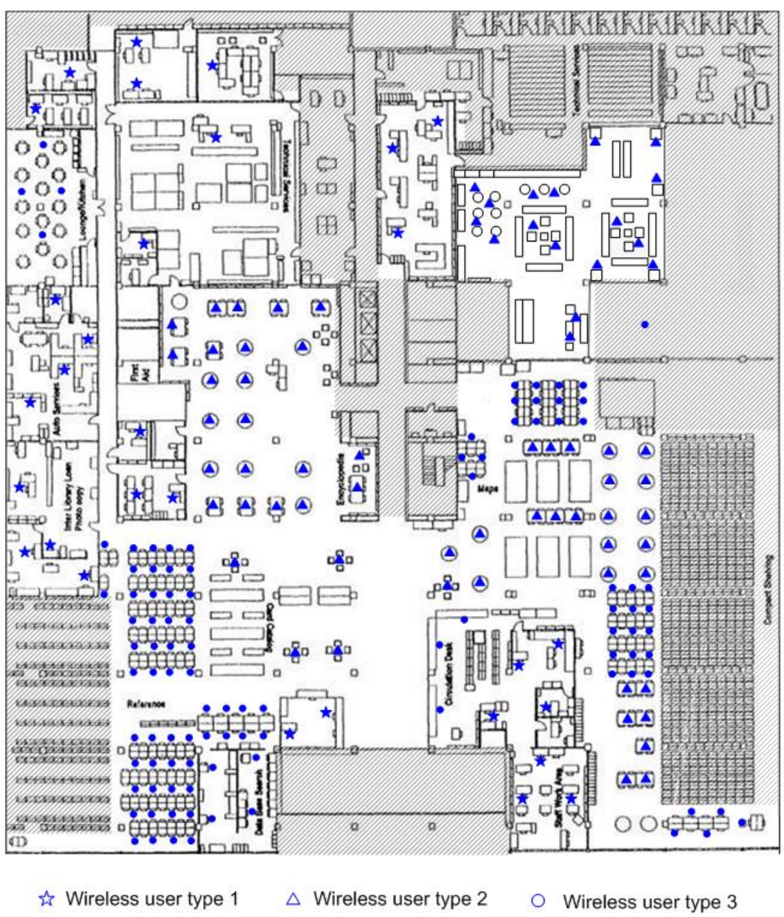
6.1.1 Experiment Settings

Numerical network design experiments were conducted for both small and large single-floor service areas. The small service area considered is the fourth floor of the School of Information Science building (SIS4) (dimension 33 × 21 meters), shown in Figure 6.1a, while the large service area is the first floor of the Hillman Library building (HL1) (dimension 66 × 75 meters), shown in Figure 6.1b. Both buildings are located on the campus of the University of Pittsburgh. SIS4 resembles a typical floor plan of an academic building containing both public and private network usage areas. The public areas are used for both scheduled and unscheduled activities. Scheduled activities occur in classrooms (rooms 403, 404, 405, 406 and 411) and laboratories (rooms 409 and 410), while unscheduled activities occur in the student lounge (room 401). The private areas of SIS4 are graduate student offices (rooms 402, 407, and 410). HL1 also contains both public and private type of usage areas. The public areas are the available study areas, and the private areas are library staff offices.

Prospective wireless users in the service area are represented by the demand node distributions shown in Figure 6.1, which are based on a preliminary site survey and information obtained from the facility staff in each location. In this figure, the symbol ● represents the demand nodes located in public areas for scheduled activities, the symbol ▲ represents the demand nodes located in public areas for unscheduled activities, and the symbol ★ represents the demand nodes located in private areas. User activity levels corresponding to each sub-area type are assumed according to network usage observations which have shown that users in private sub-areas tend to be more active in



a. The fourth floor of the School of Information Science building (SIS4)



b. The first floor of Hillman Library (HL1)

Figure 6.1 Floor plans and demand node distribution of SIS4 and HL1 data transfer activities than those in other areas, while users in the public areas of schedule-based activities tend to be more idle than those in other areas [41, 44, 45]. The average user data rates are taken from observed network usage characteristics [41, 44, 45]. The network usage characteristics are summarized in Table 6.1.

Table 6.1 Network usage characteristics

Sub-areas	User activity level	Average user data rate (Kbps)
Type 1: Private sub-areas (e.g., graduate student and library staff offices)	$\alpha_1 = 0.50$	$R_1 = 460$
Type 2: Public sub-areas for unscheduled activities (e.g., library study areas, student lounge)	$\alpha_2 = 0.40$	$R_2 = 260$
Type 3: Public sub-areas for schedule-based activities (e.g., classrooms, laboratories)	$\alpha_3 = 0.35$	$R_3 = 80$

Input parameters of the network design problem are summarized in Table 6.2. The domains of the variables have been tailored to the 802.11b specification. We employed the log distance path loss model and the omnidirectional antenna (halfwave dipole) with gain G_{AZ} of 2.5 dB to estimate the radio propagation characteristics in the service area. Other input parameters were selected based on the service environments and the building structures. This design aims for 95% coverage availability at the edge of AP coverage areas [62]. In this case, a fading margin of 5.75 dB is applied in the signal coverage calculation.

Table 6.2 Summary of parameters used in the WLAN design

Parameter	Definition	Value
Domains of Variables:		
D_p	Set of candidate power levels for variable p_j	{0, 7, 13, 15, 17, 20, 24} in dBm [87]
D_f	Set of candidate frequency channels for variable f_j	{2.412, 2.437, 2.462} in GHz [87]
D_d	Domain of binary variable d_{ij}^{t}	{0, 1}
D_g	Domain of binary variable g_{hj}	{0, 1}
$D_{(x,y)}$	Domain of variable (x_j, y_j) for $\forall j \in A$ (in meter)	For SIS4, $0 < x_j < 33$, $0 < y_j < 21$ For HL1, $0 < x_j < 66$, $0 < y_j < 75$
Static Parameters:		
α_t	User active level defines percentage of wireless users in sub-area type t that are engaged in data transfer activities (i.e., participating in channel contention and sharing AP capacity)	See Table 6.1
R_t	Average user data rate requirement in sub-area type t	
$P_{Rthreshold}$	Received sensitivity threshold	-80 dBm [88]
$SIR_{threshold}$	Signal to interference ratio threshold	10 dB [88]
C_j	Data rate capacity of the ap_j for $\forall j \in A$	11 Kbps [88]
Path loss Parameters:		
d_0	Reference distance d_0	1 meter [12]
n	Path loss exponent	3.3 [71]
δ	Standard deviation representing shadow fading	3.5 dB [72]
Antenna Parameters:		
G_{AZ}	Antenna gain (peak directivity)	2.5 dB [57]

<u>Note</u>:

- *The frequencies 2.412, 2.437, and 2.462 GHz are denoted by channels number 1, 6, and 11, respectively.*
- *The transmit power of 0, 7, 13, 15, 17, 20 and 24 dBm are denoted by power levels 0, 1, 2, 3, 4, 5, and 6, respectively.*

The service area is divided into grids of size 1m × 1m. The grid points represent possible locations for access points and specify the Signal Test Points (STPs). Figure 6.2 illustrates STPs in the SIS4 service area.

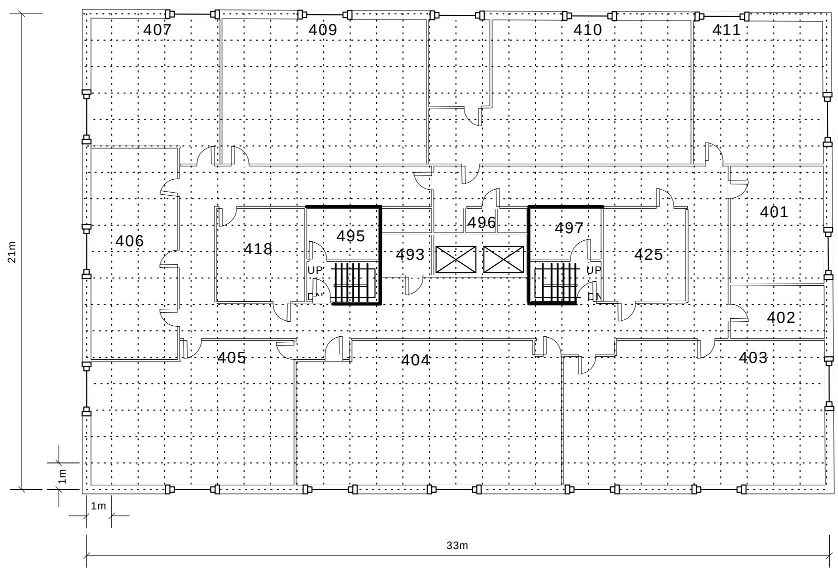


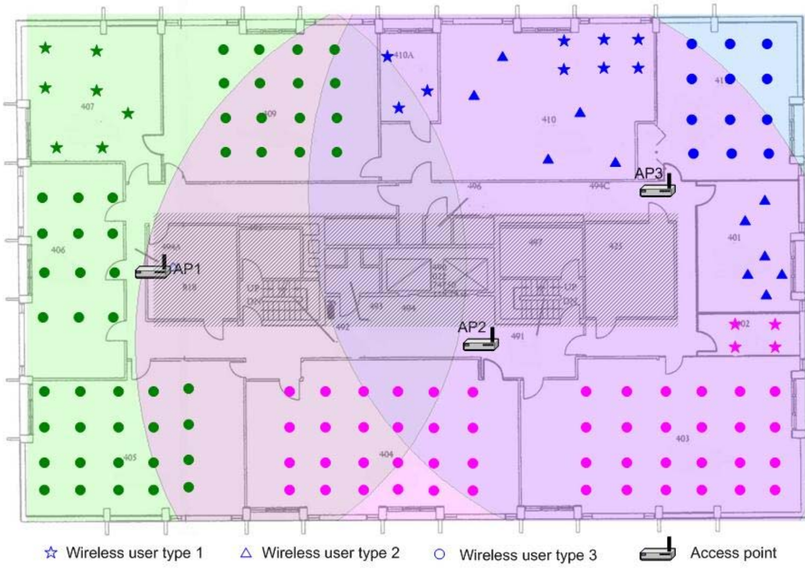
Figure 6.2 Grid point resolution of the small service area (SIS4)

6.1.2 Numerical Results

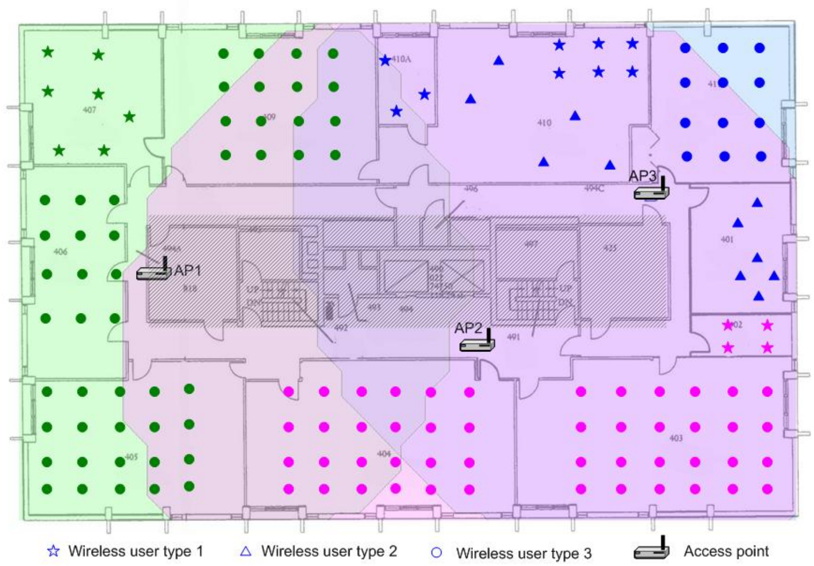
We apply the proposed demand-based WLAN design model to the SIS4 and HL1 network service areas. The heuristic solution technique described in chapter four is used to generate WLAN configurations. Figures 6.3 and 6.4 depict the resulting WLAN configurations for SIS4 and HL1, respectively. Figures 6.3a and 6.4a show contour plots of the APs' signal coverage areas, representing area in which the received signal strength from a particular AP is at least equal to the receiver sensitivity level. Such contour plots are circular in shape due to the log distance path loss model used in the path loss calculation. Figures 6.3b and 6.4b show contour plots of the APs' basic service areas, representing the area around the APs in which the SIR level is at least equal to the specified SIR threshold. Within such an area, the signal quality is good enough to allow communication between the wireless terminal and a particular AP. The shading of the contour plot figures corresponds to the frequency channels assigned to APs. In this case, channels 1, 6, and 11 are indicated by shades of green, pink, and blue, respectively. Demand nodes are colored to show which AP they are assigned. Figures 6.3c and 6.4c provide the APs' parameters, including location, frequency channel, and power level. It also provides information about the figures' colors.

For the SIS4 service area, application of the demand-based WLAN design model and the heuristic solution technique yields a network configuration using three APs with the locations, power levels and frequency channels shown in Figure 6.3. The obtained network configuration provides radio signal coverage across the service area according to the resulting signal strength and SIR evaluation at specified STPs. Based on the CSMA/CA capacity analytical models [67], the obtained network configuration can satisfy average data rate requirements to wireless users in the service area.

For the HL1 service area, the network configuration utilizes seven APs. Figure 6.4 depicts the resulting network configuration in which the APs' frequency channels, power levels and locations are assigned so that interference in the service area is avoided and the specified SIR threshold is met.



a. Contour plot of APs' signal coverage area



b. Contour plot of APs' basic service area

AP	Location	Frequency Channel	Power level	Colors represent user association to AP
1	(6.0,10.0)	1	6	
2	(19.0,7.0)	6	6	
3	(26.0,13.0)	11	6	

c. WLAN configuration

Figure 6.3 WLAN network configuration for a small service area (SIS4)

a. Contour plot of APs' signal coverage area

b. Contour plot of APs' basic service area

AP	Location	Frequency Channel	Power level	Colors represent user association to AP
1	(18.0,13.0)	11	5	
2	(15.0,43.0)	1	5	
3	(51.0,10.0)	11	4	
4	(40.0,50.0)	1	5	
5	(10.0,20.0)	6	5	
6	(16.0,52.0)	11	5	
7	(35.0,28.0)	6	6	
8	(49.0,26.0)	1	4	

c. WLAN configuration

Figure 6.4 WLAN network configuration for a large service area (HL1)

6.2 Multi-floor WLAN design

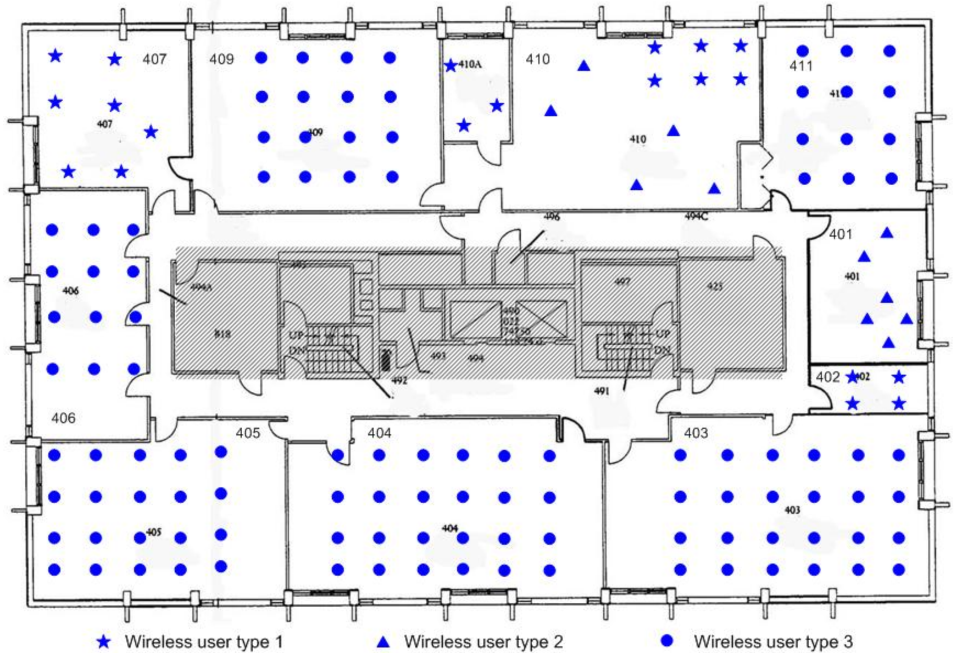
A WLAN's service area may include some rooms on a single floor, an entire floor or even multiple floors. In a WLAN covering multiple floors, interference caused by nearby APs operating at co-channels or overlapping channels is an important concern during the network design phase. APs located on adjacent floors may interfere with each other if the design does not coordinate the AP placement, frequency assignment, and power level assignment.

In this section, we consider a multi-floor WLAN design using the proposed demand-based design model and heuristic solution technique to develop an efficient network configuration.

6.2.1 Experimental Settings

In this study, we consider the design of a service area aiming to cover the fourth and fifth floors (SIS4 and SIS5) of the School of Information Science building as shown in Figure 6.5. SIS4 is a typical floor plan of an academic building containing a large central structure with elevator shafts, stairwells, and restrooms. Around this center, partitioned by office walls, are classrooms, graduate student offices and a student lounge. SIS5 also has a large central structure with the same components as that of SIS4. Around this are administrative offices, conference rooms, and classrooms. The expected wireless users on these two floors are represented by the demand node distribution map shown in Figure 6.5. Wireless users located in private office spaces can be identified in advance while prospective users in classrooms and laboratories can be estimated from the

a. 4th floor



b. 5th floor

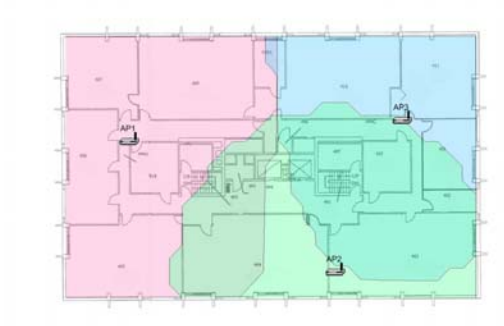
Figure 6.5 Floor plans of the fourth and fifth floors of the School of Information Science building

available seats in the room. For example, the number of wireless users in the classrooms is limited by the number of seats provided. The network usage characteristics of prospective wireless users used in this section are the same as those shown in Table 6.1 (section 6.1).

6.2.2 Numerical Results

The demand-based design model and the developed heuristic solution technique are applied to solve this multi-floor WLAN design problem. The resulting network utilizes five APs. Based on traffic demand, three of the APs (AP1, AP2, and AP3) are located on the fourth floor. Two APs (AP4 and AP5) are located on the fifth floor. Figure 6.6 depicts signal coverage on each floor. Figures 6.6a and 6.6b show signal coverage on the floor where the APs are located, with 6.6a showing coverage on the fourth floor and 6.6b showing coverage on the fifth floor. The basic service area (BSA) of each AP is shaded in a different color, and each color is associated with the frequency channel assigned to that AP. Green, blue and pink represent the BSAs of APs operating at channels 1, 6 and 11, respectively. Figure 6.6c presents signal coverage from two APs located on the fifth floor to the area of the fourth floor while Figure 6.6d presents signal coverage from three APs located on the fourth floor to the area on the fifth floor.

The process of assigning frequency channels in the heuristic solution technique coordinates the channels assigned to APs on different floors to limit interference in the network. Figure 6.7 depicts overall signal coverage on each floor. The signal penetrating from each floor is superimposed on the contour plot of the other floor. The demand node assignment is represented by different colors as listed in Figure 6.7c. For example, the

<table>
<tr><th>4th floor</th><th>5th floor</th></tr>
<tr><td>

a. Signal on 4th floor from
AP1, AP2, AP3 (located on 4th floor)</td><td>

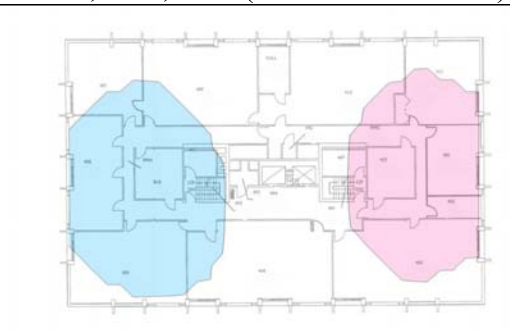
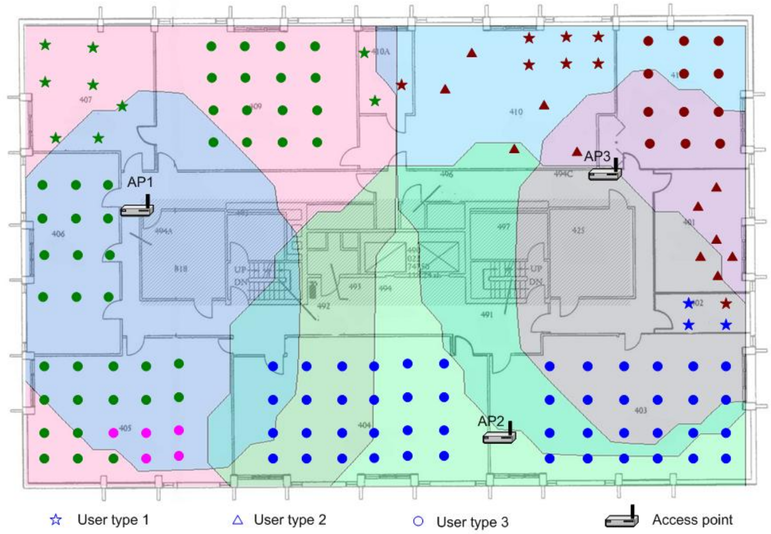
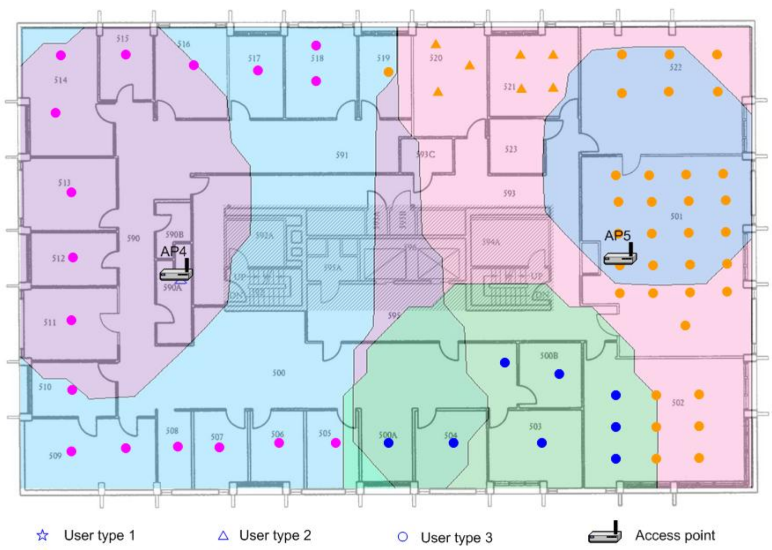
b. Signal on 5th floor from
AP4, AP5 (located on 5th floor)</td></tr>
<tr><td>
c. Signal on 4th floor from
AP4, AP5 (located on 5th floor)</td><td>
d. Signal on 5th floor from
AP1, AP2, AP3 (located on 4th floor)</td></tr>
</table>

Figure 6.6 Signal coverage from APs located on different floors

demand node assigned to AP1 is green, and the demand node assigned to AP2 is blue. We can see that the blue demand nodes on the fifth floor are assigned to AP2 on the fourth floor and that the pink demand nodes on the fourth floor are assigned to AP4 on the fifth floor. This demonstrates that in this WLAN design scenario, the APs can handle traffic generated by wireless users on different floors.



a. Overall signal coverage and demand node association to APs on the 4th floor



b. Overall signal coverage and demand node association to APs on the 5th floor

AP	Location	Floor number	Frequency channel	Power level	Colors represent user association to AP
1	(5.0,12.0)	4	11	5	
2	(21.0,2.0)	4	1	6	
3	(26.0,13.0)	4	6	5	
4	(7.0,9.0)	5	6	6	
5	(26.0,10.0)	5	11	6	

c. WLAN configuration

Figure 6.7 WLAN configuration of the 4th and 5th floors of the SIS building

6.3 Extend Indoor WLAN Design to a Courtyard Area

As WLANs become increasingly popular, the need of their services expands from inside buildings to include outside areas such as courtyards [2]. In WLANs in which the target service area consists of both indoor and outdoor areas, signal penetration from outdoor APs may affect the deployment of indoor APs [89-91]. The WLAN design process needs to coordinate the assignment of parameters for both the indoor and outdoor APs. In this section, we illustrate the application of the demand-based WLAN design model and the heuristic solution technique developed in this dissertation to a design scenario that extends the service coverage to include a courtyard area.

6.3.1 Experiment Settings

We consider the design of a service environment which extends from inside Hillman Library to a courtyard area outside. The layout of the service area and the estimated demand node distribution are shown in Figure 6.8. We consider network usage characteristics as described in Table 6.1.

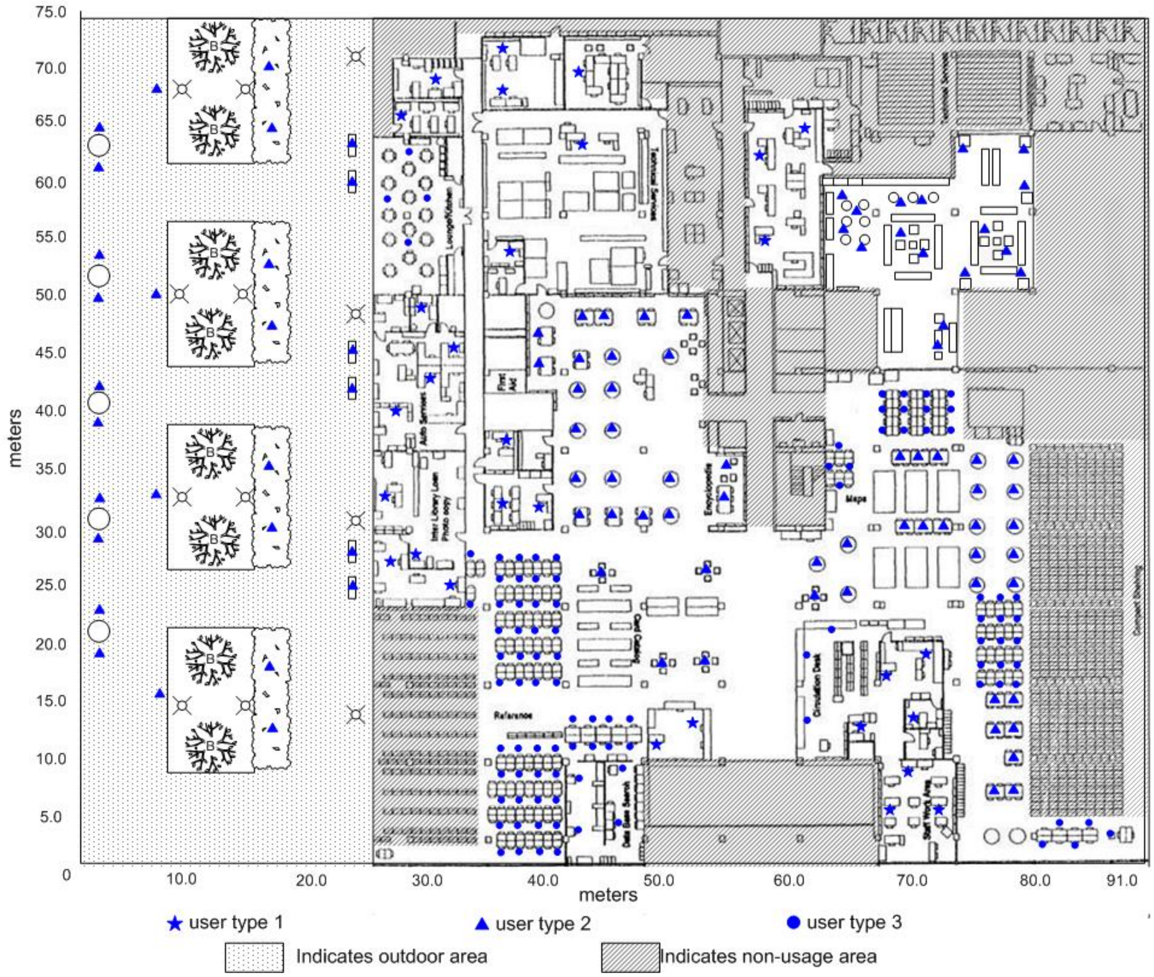


Figure 6.8 Layout of service area and demand node distribution of Hillman Library and the courtyard area next to the building

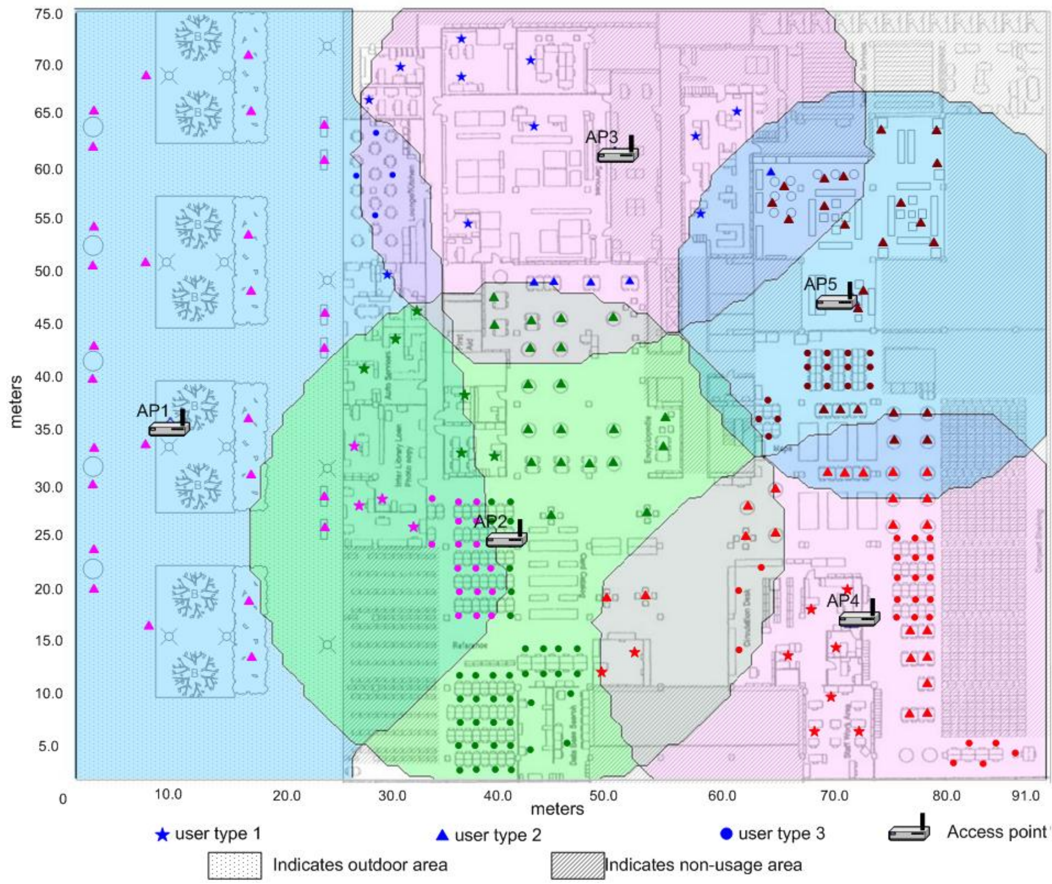
The demand-based design model imposes network requirements upon the indoor and outdoor service area. The heuristic solution technique is applied to solve the WLAN design problem. The empirical outdoor path loss model for 802.11b developed for the university campus environment using a transmitter height of 4 meters (as written in equation 6.1) [92] is applied in the path loss calculation of the outdoor area while the indoor log distance path loss model [12] (as shown in equation 2.1) and the parameters for indoor radio propagation are applied to the path loss calculation of the indoor area. The aggregate penetration loss of signal from outdoor to indoor area is also considered. An attenuation of 20 dB is added to the path loss of signal penetration from outdoor to indoor [91].

$$\overline{PL(d)} = PL(d_0) + 10n_f \, log_{10}\left(\frac{d}{d_0}\right) + 5 + 0.03d \qquad (6.1)$$

where $\overline{PL(d)}$ is the average path loss (dB), $PL(d_0)$ is a fixed path loss at a reference distance d_0, n_f is a path loss exponent for free space, and d is the distance between the transmitter and receiver in meters. In these tests, the values of d_0 = 1 m. and n_f = 2 are used.

6.3.2 Numerical Results

The heuristic solution technique can harmonize the placement of APs and the assignment of frequency channels of indoor and outdoor APs. Figure 6.9 depicts the resulting network configuration, in which five APs provide service coverage in the target area. AP1 is located on a lamppost in the courtyard, and the other four APs are located within the library. AP1 is assigned frequency channel 6, and its adjacent APs (i.e., AP2 and AP3) are assigned frequency channels 1 and 11, respectively. According to the path loss calculation used in the design, a signal from AP1 will penetrate into the building and carry some of the traffic load from demand nodes in the building (in Figure 6.9, such demand nodes are colored in pink).



a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(8.8,34.0)	6	6	
2	(41.0,23.0)	1	6	
3	(50.0,60.0)	11	6	
4	(72.0,15.0)	11	6	
5	(72.0,46.0)	6	6	

b. WLAN configuration

Figure 6.9 WLAN configuration for Hillman Library and the courtyard next to the building

6.4 Effects of User Activity Level

The user activity level is defined as the percentage of prospective wireless users who associate to the APs in the service area and use the network (e.g., participate in the wireless channel contention and share AP capacity) at a particular time. The user activity level varies across different areas with different usages or functionalities. As reported in network trace studies [41, 42, 44, 45], WLAN usage pattern and user behavior are correlated with users' locations. For example, the amount of traffic generated in areas such as auditoriums and classrooms in which a high number of users associate to the APs, is quite low compared to the traffic generated by users in private offices where user activity is high [41, 42, 44, 45]. We account for WLAN usage patterns and user behaviors in the demand-based WLAN design model through the use of a user activity level parameter. This parameter indicates the activity of prospective wireless users on the network and the expected amount of traffic that needs to be carried in the network.

In this section, we aim to investigate the effects of the user activity level on the network configuration and the performance of the WLAN obtained from the proposed demand-based WLAN design model.

6.4.1 Experiment Settings

Two types of service environments are studied. They include the fourth floor of the School of Information Science building (SIS4) and the first floor of Hillman Library (HL1), shown Figure 6.1 (section 6.1). For these two service environments, we study the WLAN design in six scenarios. Scenarios 1 through 3 consider the SIS4 environment, while scenarios 4 through 6 consider the HL1 environment. We vary the user activity level (α_t) so that scenarios 1 and 4 reflect an area with low network traffic, scenarios 2 and 5 replicate an area with medium network traffic, and scenarios 3 and 6 reflect an area with high network traffic.

Note that user activity levels for the medium traffic case are based on the WLAN usage patterns and characteristics experienced and reported in the literature in which users in private sub-areas tend to be more active in data transfer activities than those located in other areas, while users in public sub-areas used for schedule-based activities tend to be more idle than users in other areas [41, 42]. The parameters of the low and high traffic cases are defined as being 40% lower and higher than those of the medium traffic case. In all scenarios, we consider the same set of average user data rates (R_t in Kbps), drawn from the network trace results reported in [40, 44]. In the private spaces, the average user data rate is 460 Kbps. In those public spaces with scheduled activities, the average user data rate is 80 Kbps. In the public spaces with unscheduled activities, the average user data rate is 260 Kbps. Parameters used in each scenario are summarized in Table 6.3.

The proposed demand-based WLAN design model was applied to the six network design scenarios of the SIS4 and HL1 environments, as described above. The heuristic solution techniques described in chapter 4 were used to generate WLAN configurations for all design scenarios. The static parameters and the domain of the variables used in the solution technique are the same as those specified in section 6.1.

Table 6.3 Parameters used in sensitivity analysis to study the effects of user activity level

Scenario	Service environment	α_1	α_2	α_3	R_1	R_2	R_3
1	SIS4, low user activity level	0.30	0.25	0.20	460	260	80
2	SIS4, medium user activity level	0.50	0.40	0.35	460	260	80
3	SIS4, high user activity level	0.70	0.55	0.50	460	260	80
4	HL1, low user activity level	0.30	0.25	0.20	460	260	80
5	HL1, medium user activity level	0.50	0.40	0.35	460	260	80
6	HL1, high user activity level	0.70	0.55	0.50	460	260	80

Note: α_t = user activity level of the prospective wireless users that are present in a service area of type t
R_t = average user data rate requirement (in Kbps) of the prospective wireless users that are present in a service area of type t

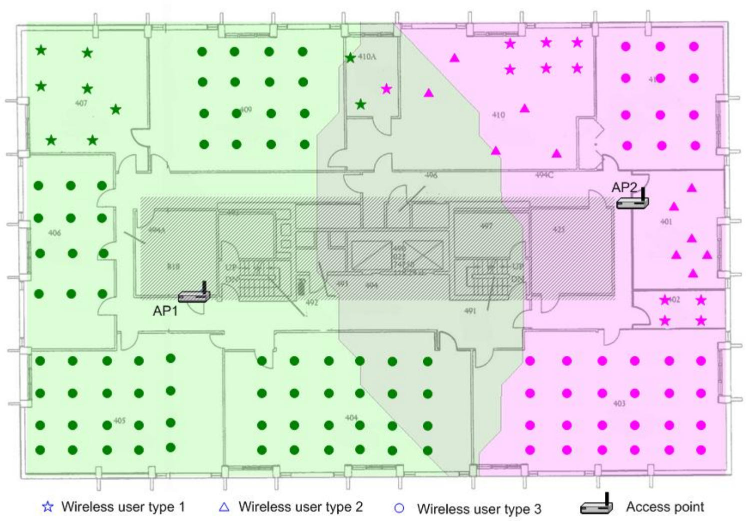
where t = 1 indicates private spaces (e.g., graduate student offices, library staff offices)
t = 2 indicates public spaces for unscheduled activities (e.g., library study areas, student lounge)
t = 3 indicates public spaces for scheduled activities (e.g. classrooms, meeting rooms)

6.4.2 Numerical Results and Discussion

6.4.2.1 Effects of User Activity Level on Network Configurations

Figures 6.10 through 6.15 shows WLAN configurations obtained for design scenarios 1 – 6, respectively. The figures show the contour plots of the APs' basic service areas, indicating regions around the APs where signal strength and SIR level are good enough to allow data communication. Colors of the APs' basic service areas correspond to frequency channel assignments. For example, in Figure 6.12, the basic service areas of APs using channels 1, 6, and 11 are shaded in blue, green, and pink, respectively. Demand nodes are colored to show to which AP they are assigned. The figures also provide details about the APs' parameters, including locations, frequency channels, and power levels. It also provides information about the colors representing user association to APs.

In the case of SIS4 with a low user activity level (scenario 1), the obtained network configuration utilizes two APs to provide radio signal coverage and accommodate the expected data traffic in the area (Figure 6.10). As the prospective wireless users become more active in data transfer activities (i.e., higher value of the user activity level), the amount of traffic aggregating at the APs and competing for the wireless channel increases. As a result, more APs are needed to support network activity. In the medium network traffic case (scenario 2), three APs are used (Figure 6.11) while in the high network traffic case (scenario 3), four APs are used (Figure 6.12).

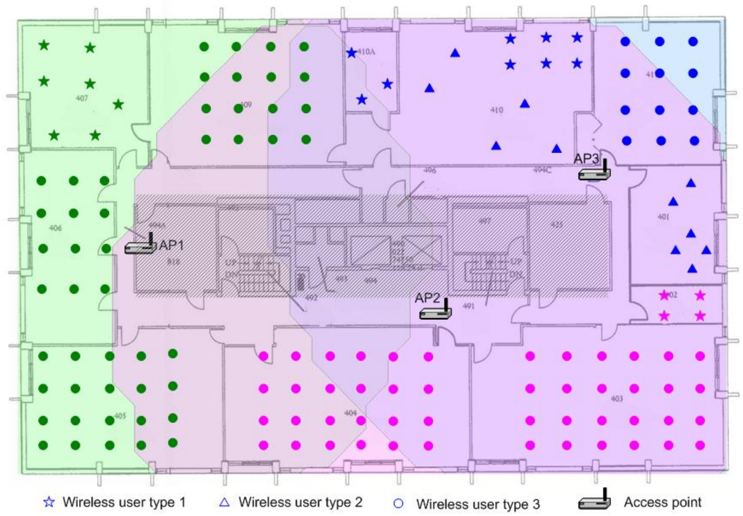


a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(8.0,8.0)	1	6	
2	(28.0,12.0)	6	6	

b. WLAN configuration

Figure 6.10 Basic service areas and WLAN configuration of SIS4, low user activity level (scenario 1)

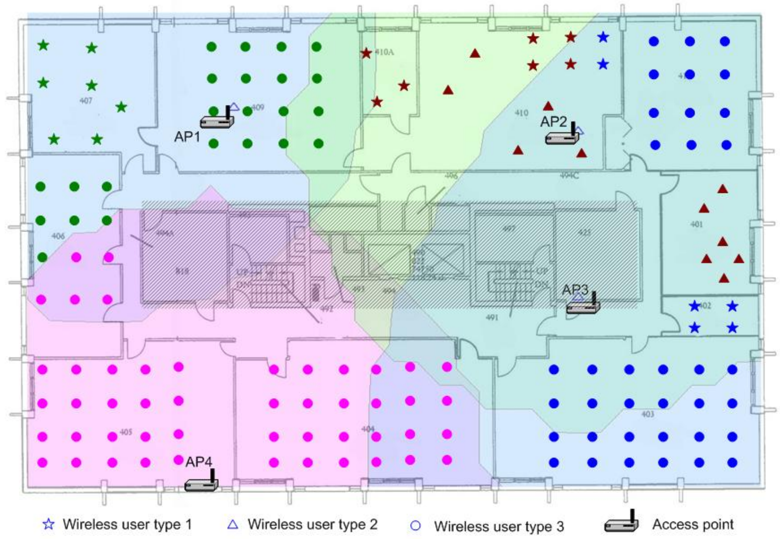


a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(6.0,10.0)	1	6	
2	(19.0,7.0)	6	6	
3	(26.0,13.0)	11	6	

b. WLAN configuration

Figure 6.11 Basic service area and WLAN configuration of SIS4, medium user activity level (scenario 2)



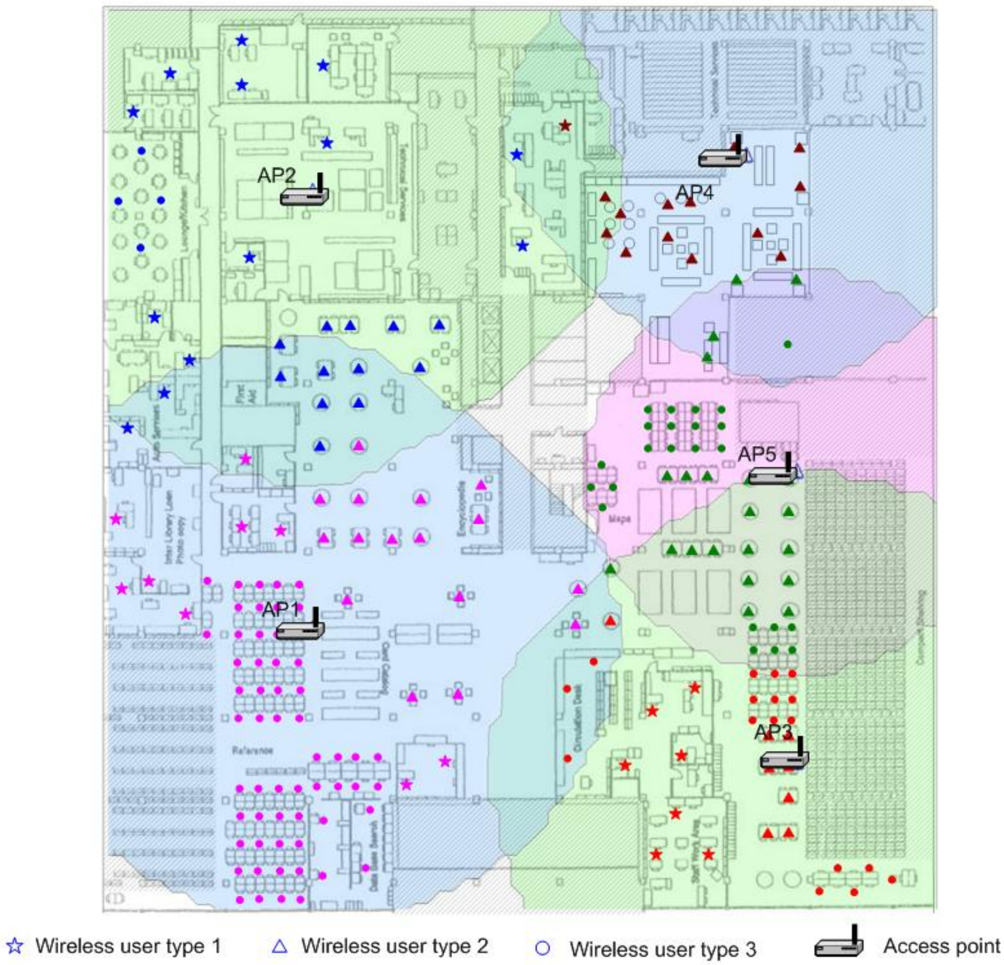
a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(9.0,16.0)	1	5	
2	(23.0,15.0)	6	5	
3	(24.0,8.0)	1	6	
4	(8.0,0.0)	11	5	

b. WLAN configuration

Figure 6.12 Basic service areas and WLAN configuration of SIS4, high user activity level (scenario 3)

The resulting WLAN system configurations for HL1 scenarios 4 through 6 are shown in Figures 6.13 – 6.15, respectively. As in the case of SIS4, more APs are used in the service areas where high traffic demand is expected. In the case of low traffic demand (scenario 4), Figure 6.13 shows a WLAN configuration using five APs to provide radio signal coverage and accommodate data traffic generated by users in the service area. In the medium network traffic case (scenario 5), Figure 6.14 shows a WLAN configuration in which seven APs are used. Finally, in the high network traffic case (scenario 6), Figure 6.15 shows the resulting WLAN configuration in which eight APs are used. As the number of APs used in the network configuration increases, the incidence of co-channel frequency reuse increases. This causes more interference between co-channel APs. As a result, the APs' basic service areas become irregularly shaped as can be observed in Figures 6.14 and 6.15.

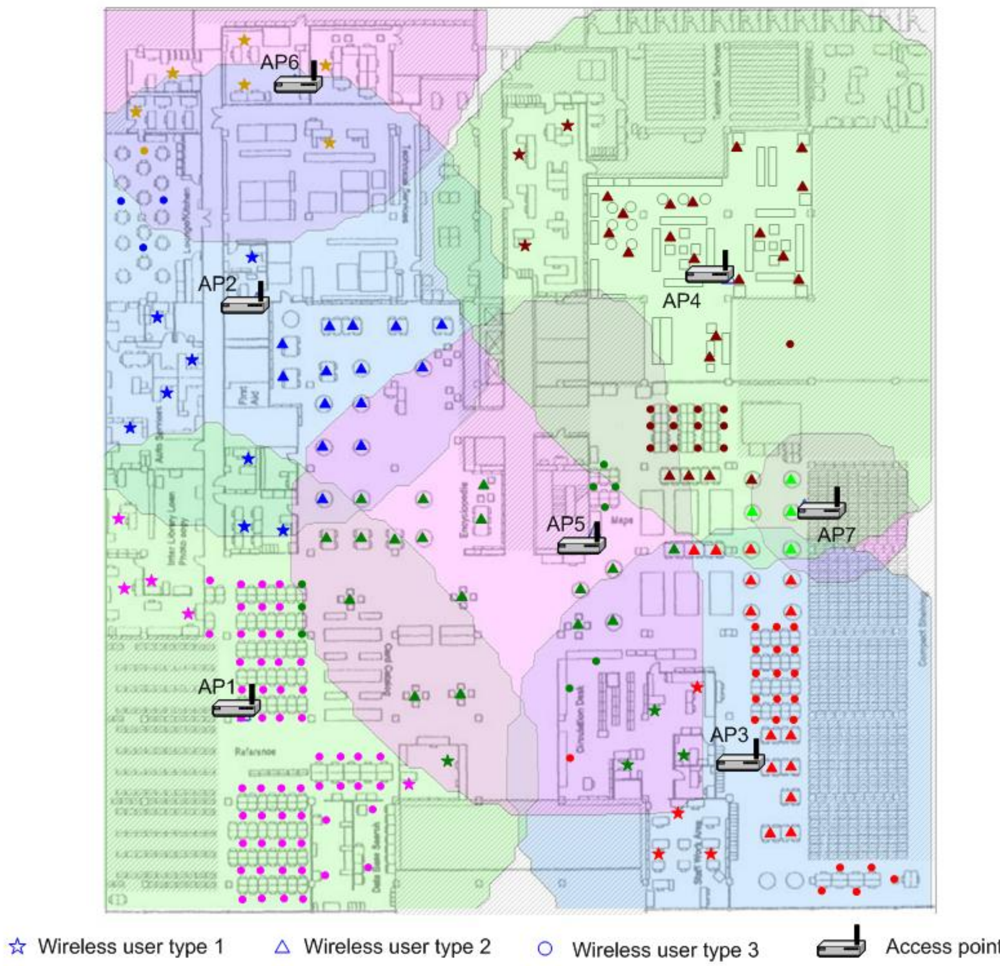


a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(16.0,23.0)	6	6	
2	(16.0,59.0)	1	6	
3	(54.0,12.0)	1	6	
4	(50.0,62.0)	6	5	
5	(54.0,36.0)	11	4	

b. WLAN configuration

Figure 6.13 Basic service areas and WLAN configuration of HL1, low user activity level (scenario 4)

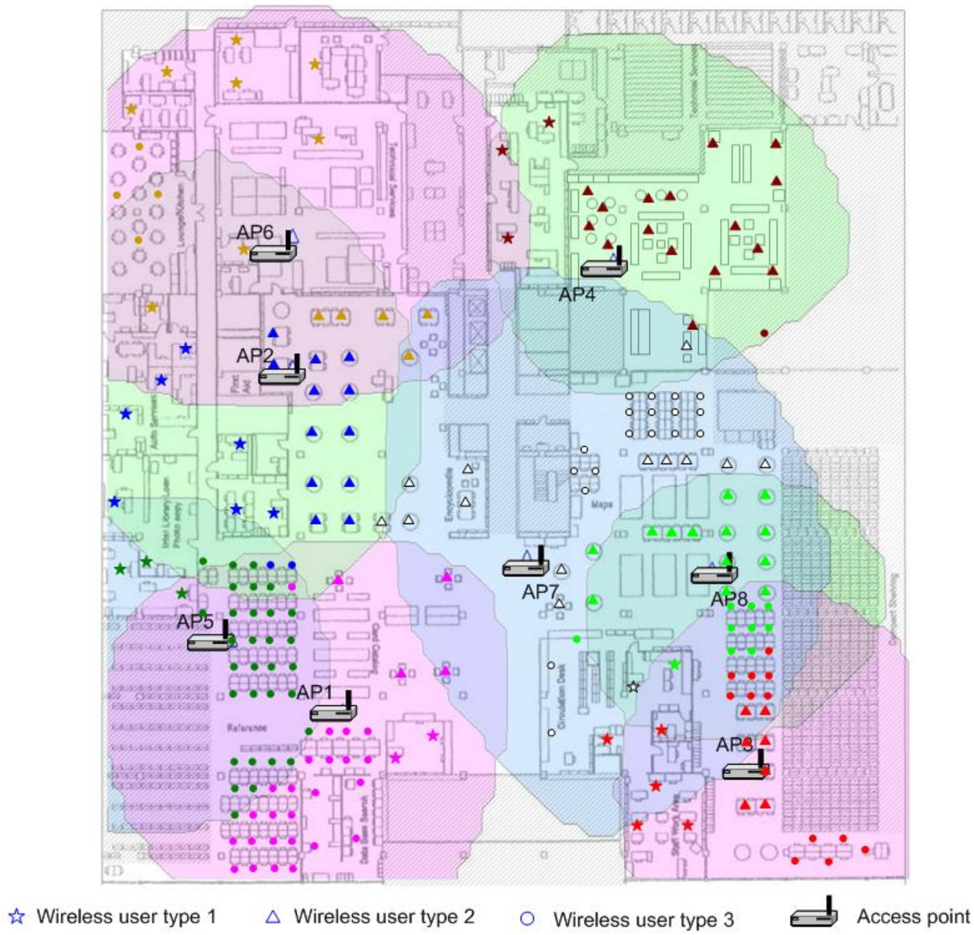


a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(11.0,16.0)	1	6	
2	(12.0,50.0)	6	5	
3	(51.0,12.0)	6	5	
4	(49.0,52.0)	1	6	
5	(38.0,31.0)	11	6	
6	(15.0,68.0)	11	4	
7	(55.0,33.0)	11	3	

b. WLAN configuration

Figure 6.14 Basic service areas and WLAN configuration of HL1, medium user activity level (scenario 5)



a. Contour plot of APs' basic service areas

AP	Location	Frequency channel	Power level	Colors represent user association to AP
1	(18.0,13.0)	11	5	
2	(15.0,43.0)	1	5	
3	(51.0,10.0)	11	4	
4	(40.0,50.0)	1	5	
5	(10.0,20.0)	6	5	
6	(16.0,52.0)	11	5	
7	(35.0,28.0)	6	6	
8	(49.0,26.0)	1	4	

b. WLAN configuration

Figure 6.15 Basic service areas and WLAN configuration of HL1, high user activity level (scenario 6)

6.4.2.2 Comparative Performance Evaluation: Demand-based WLAN Design vs. Coverage-based WLAN Design

Coverage-based WLAN design is the traditional and existing WLAN design approach and it focuses on providing radio signal coverage in the service region [24, 28, 30-32]. In such work, the network traffic generated by wireless users in the service area is not considered. To demonstrate the superiority of the proposed demand-based WLAN design over the existing coverage-based design, we derive the comparative performance evaluation, considering the WLAN scenarios 1 through 6 described previously.

The network configuration shown in Figure 6.10, designed for the case of a low network traffic (scenario 1), is considered to be the coverage-based WLAN configuration for the SIS4. The network configuration shown in Figure 6.13, designed for the case of a low network traffic (scenario 4), is considered to be the coverage-based WLAN configuration for the HL1. The reason is that such network configurations (Figures 6.10 and 6.13) employ the minimum number of APs that can provide the required received signal strength and interference level in the service region according to the log-distance path loss model used in the path loss calculation of the design process. Figures 6.11 and 6.12 show the demand-based WLAN configurations designed for SIS4 scenarios 2 and 3, respectively. Figures 6.14 and 6.15 show the demand-based WLAN configurations designed for HL1 scenarios 5 and 6, respectively.

Tables 6.4 and 6.5 compare the network performance of the demand-based WLAN configurations against that of the coverage-based WLAN configurations under different network traffic conditions for SIS4 and HL1, respectively. Performance matrices used for the comparison include:

- Lowest average data rate available to an individual user present in a particular usage area type (L_{avg}, Kbps)
- Percentage of prospective wireless users in the service area that can obtain their average data rate as required (% Satisfaction)
- The number of APs used in the obtained network configurations
- Average AP throughput, comparing all APs used in the service area (AP_{avg}, Kbps)

Table 6.4 presents a comparative performance evaluation of SIS4 in three network traffic scenarios (scenario 1 through 3). Performances of the demand-based WLAN configurations are compared to that of the coverage-based WLAN configurations based on the performance matrices described above. In scenario 1 (low network traffic scenario), the coverage-based WLAN configuration with 2 APs can provide the average user data rate as required. The network traffic increases in scenarios 2 and 3, and the demand-based WLAN configurations (employing 3 and 4 APs, respectively) outperform the coverage-based WLAN configurations according to the resulting average data rate available to wireless users in the network. In Table 6.4, the lowest average obtainable user data rates are compared to the average user data rate requirement. The demand-based WLAN configuration can meet the average user data rate requirement in all scenarios considered. The lowest average data rates available to users meets the average data rate requirements, providing 100% of prospective wireless users in the service area with the average data rate as required. The coverage-based WLAN configuration, on the other hand, cannot provide the required average data rate to users in the service area of the scenarios 2 and 3. As the user activity levels increase in scenarios 2 and 3, the number of active users that are simultaneously communicating and sharing the APs' capacity increases. According to the capacity analytical model, as the number of users simultaneously connecting to the AP increases, the effective AP capacity decreases. As a result, the average obtainable user data rate becomes lower than if fewer users were sharing the same AP. Table 6.4, looking at scenario 2 (medium network traffic), shows that only 8% of wireless users can obtain the average data rate as required while 92% of the users suffer as a result of competing for the data channel and sharing APs under high network traffic conditions; the result is a lower average data rate than required. In scenario 3 (high network traffic), none of the wireless users in the area can obtain the required data rate because all APs in the network are overloaded with high traffic volume.

As shown in Table 6.5, similar conclusions can be made for the HL1 service area. The coverage-based WLAN configuration using 5 APs can provide the average user data rate required in the low network traffic case (scenario 4). However, when faced with the higher user activity levels of scenarios 5 and 6, AP1 and AP5 of the coverage-based WLAN configuration suffer from higher traffic load, resulting in lower average AP throughput compared to that of the scenario 4. The percentage of wireless users who can obtain average data rate as required falls to 60% and 30% in scenario 5 and 6, respectively.

Table 6.4 Comparative performance evaluation of WLAN configurations for SIS4

		Low network traffic (scenario 1)			Medium network traffic (scenario 2)			High network traffic (scenario 3)		
		t=1	t=2	t=3	t=1	t=2	t=3	t=1	t=2	t=3
Lowest average data rate available to an individual use located in a particular usage area type (Kbps)	Average data rate required	460.0	260.0	80.0	460.0	260.0	80.0	460.0	260.0	80.0
	Demand – based design	737.2	491.5	98.3	608.4	405.6	81.2	602.1	401.4	80.3
	Coverage - based design	737.2	491.5	98.3	457.5	260.0	52.0	257.2	171.5	34.3
% wireless users in service area obtaining average data rate as required	Demand – based design	100%			100%			100%		
	Coverage - based design	100%			8%			0%		
# APs used in the obtained network configuration	Demand – based design	2			3			4		
	Coverage - based design	2			2			2		
Average AP throughput of each AP used in the service area (Mbps)	Demand – based design	ap_1=4.100, ap_2=4.669			ap_1=3.970, ap_2=5.232, ap_3=3.756			ap_1=4.215, ap_2=5.323, ap_3=4.215, ap_4=3.291		
	Coverage - based design	ap_1=4.100, ap_2=4.669			ap_1=3.633, ap_2=4.130			ap_1=3.360, ap_2=3.823		

Note: *t* = 1 indicates private spaces (e.g., graduate student offices, library staff offices)
t = 2 indicates public spaces for unscheduled activities (e.g., library study areas, student lounge)
t = 3 indicates public spaces for scheduled activities (e.g., classrooms, meeting rooms)

Table 6.5 Comparative performance evaluation of WLAN configurations for HL1

		Low network traffic (scenario 4)			Medium network traffic (scenario 5)			High network traffic (scenario 6)		
		t=1	t=2	t=3	t=1	t=2	t=3	t=1	t=2	t=3
Lowest average data rate available to an individual user located in a particular usage area type (Kbps)	Average data rate required	460.0	260.0	80.0	460.0	260.0	80.0	460.0	260.0	80.0
	Demand – based design	609.8	406.6	81.3	605.8	403.8	80.8	618.0	412.0	82.4
	Coverage - based design	609.8	406.6	81.3	328.2	218.8	43.8	207.7	138.5	27.7
% wireless users in service area obtaining average data rate as required	Demand – based design	100%			100%			100%		
	Coverage - based design	100%			60%			30%		
# APs used in the obtained network configuration	Demand – based design	5			7			8		
	Coverage - based design	5			5			5		
Average AP throughput of each AP used in the service area (Mbps)	Demand – based design	ap_1=4.431, ap_2=6.646, ap_3=5.805, ap_4=7.166, ap_5=5.184			ap_1=3.450, ap_2=5.966, ap_3=4.868, ap_4=5.250, ap_5=5.729, ap_6=7.590, ap_7=7.084			ap_1=4.140, ap_2=6.058, ap_3=5.323, ap_4=6.152, ap_5=3.776, ap_6=6.254, ap_7=4.656, ap_8=5.257		
	Coverage - based design	ap_1=4.431, ap_2=6.646, ap_3=5.805, ap_4=7.166, ap_5=5.184			ap_1=3.873, ap_2=6.023, ap_3=5.322, ap_4=6.812, ap_5=4.573			ap_1=3.642, ap_2=5.730, ap_3=4.887, ap_4=6.090, ap_5=4,257		

6.5 Effects of Shadow Fading

The signal strength received at a particular location will vary due to fluctuations in the signal caused by "shadow fading," experienced as the signal goes through obstructions. Shadow fading can be characterized as a lognormal distribution where the fading component has a Gaussian distribution with a zero mean and a particular standard deviation determined based upon the environment and the surroundings [12]. As a result, the received signal strength at some locations in a service area may fall below a desired receiver sensitivity threshold. In order to provide adequate service coverage to the target area, the transmitted power needs to be increased beyond the level required to overcome the fluctuation in the received signal strength due to shadow fading effects. The amount by which the transmitted power needs to be raised is called the fading margin.

The fading margin is determined based upon the desired percentage of coverage availability at the boundary of the AP's coverage area (cell) [12, 71]. For example, we want to provide 95% coverage availability at the edge of the cell of a particular network in which the standard deviation (σ) of the shadow fading component is 4 dB. To find a fading margin (F) for this case, we solve the equation $\frac{1}{2} erfc\left(\frac{F}{\sigma\sqrt{2}}\right) = 0.05$ by using Matlab and obtain F = 6.58 dB.

In this study, the effects of shadow fading on the WLAN configurations obtained from the design process are observed.

6.5.1 Experimental Settings

This study considers WLAN design for the first floor of Hillman Library (HL1), dimension 66m. × 75m., at the University of Pittsburgh, as shown in Figure 6.1b (section 6.1). The demand node distribution and the network usage characteristics are the same as that described in section 6.1.

In planning a wireless network, designers usually aim to provide 95-99% coverage availability at the edge of the cell [62]. In this experiment, our design aims for 95% coverage availability at the edge of the cell. We consider five values of the shadow fading. The corresponding fading margin for each value of the shadow fading is presented in Table 6.6. These fading margins have been applied to the link budget calculation of the design process. The effects of shadow fading on the resulting WLAN configuration are observed.

Table 6.6 Shadow fading and corresponding fading margins

Shadow fading (dB)	Fading margin (dB) for 95% coverage availability at the edge of cells
1	1.645
2	3.290
3	4.935
4	6.580
5	8.225

6.5.2 Numerical Results and Discussions

The initial network configuration used to derive the WLAN configuration for the case of shadow fading of 1 dB is obtained from the construction phase and the initial frequency channel assignment phase. Table 6.7 shows the parameters of this initial network configuration. The WLAN configuration resulting from consideration of shadow fading of 1 dB is called WLAN1. We aimed to observe how the network configurations change from WLAN1 as the shadow fading increases. So, WLAN1 is then used as the initial configuration to determine WLAN configurations for the remaining shadow fading cases. Table 6.8 summarizes the resulting WLAN configurations obtained for each shadow fading value considered in this experiment. In Table 6.8, parameters of APs in the configurations altered from that of the WLAN1 are highlighted to indicate differences between network configurations.

Table 6.8 shows that the WLAN configurations obtained for each case of shadow fading are slightly different from that of WLAN1. As the shadow fading increases, power levels of some APs need to be increased to compensate for higher signal reduction, and locations of some APs are adjusted to prevent excessive interference between co-channel APs due to higher signal fluctuation. Figure 6.16 shows the locations of APs with all resulting WLAN configurations superimposed on each other. We can see that APs' locations do not change much in the resulting network configurations when different values of shadow fading are considered.

Table 6.7 Initial network configuration

Initial network configurations	AP0	AP1	AP2	AP3	AP4	Ap5	AP6
Location (x,y)	(15,14)	(12,50)	(51,12)	(49,55)	(25,30)	(18,69)	(52,30)
Frequency channel	11	6	1	11	6	1	1
Power level	4	4	4	6	4	4	4

Table 6.8 Resulting WLAN configurations

Shadow fading (dB)	Resulting WLAN configurations	AP0	AP1	AP2	AP3	AP4	AP5	AP6
1 (WLAN1)	Location (x,y)	(17,20)	(12,50)	(56,6)	(46,44)	(15,12)	(9,74)	(52,33)
	Freq. channel	6	1	11	11	11	6	1
	Power level	6	4	4	6	5	3	3
2	Location (x,y)	(17,20)	(12,50)	(56,6)	(46,44)	(15,12)	(9,74)	(52,33)
	Freq. channel	6	1	11	11	11	6	1
	Power level	6	4	5	6	5	3	3
3	Location (x,y)	(17,20)	(12,50)	(54,7)	(46,44)	(15,12)	(11,72)	(52,33)
	Freq. channel	6	1	11	11	11	6	1
	Power level	6	4	5	6	5	3	3
4	Location (x,y)	(17,20)	(12,50)	(51,10)	(46,44)	(15,12)	(12,71)	(52,33)
	Freq. channel	6	1	11	11	11	6	1
	Power level	6	4	5	6	5	3	3
5	Location (x,y)	(17,20)	(12,50)	(51,11)	(48,55)	(15,12)	(9,74)	(52,33)
	Freq. channel	6	1	11	11	11	6	1
	Power level	6	5	5	6	5	4	4

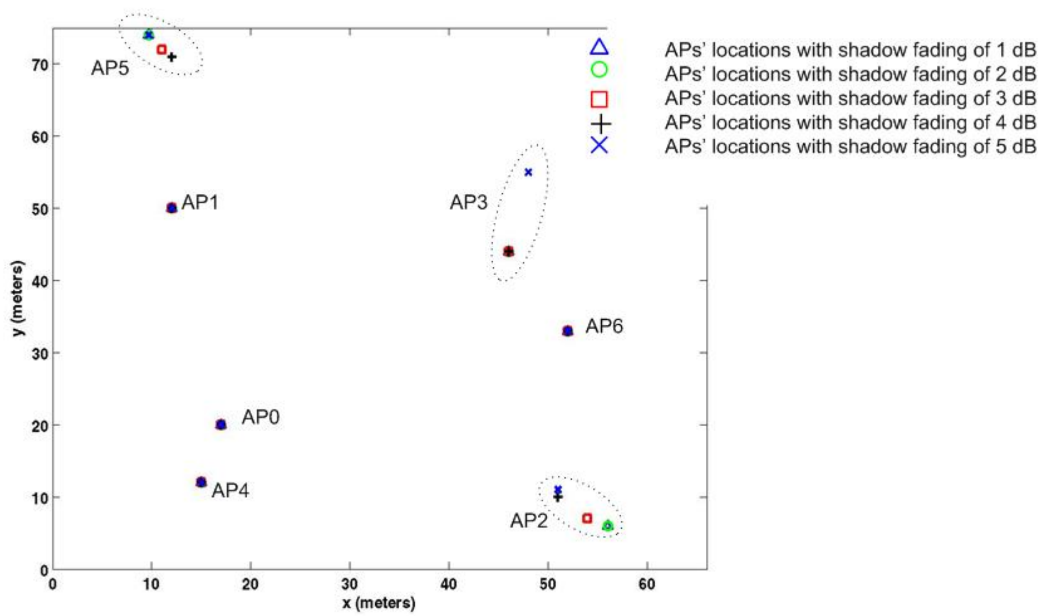


Figure 6.16 Locations of APs in WLAN configurations obtained for different values of shadow fading.

6.6 Effects of Path Loss Models used in the WLAN Design

In the WLAN design, signal strength calculation is a major task to estimate the coverage area of the APs. Path loss models are used to calculate this signal coverage. It relates the loss of signal strength to distance and environment between the AP and the receiver. Several path loss models have been developed for use in indoor wireless network design as reported by Rappaport [12]. In this study, the effects of using different path loss models on WLAN configurations are investigated.

6.6.1 Experiment Settings

We conducted this study by considering the design of two service areas, the fourth floor of the School of Information Science building (SIS4), dimension 33m × 21m (Figure 6.1a) and the first floor of Hillman Library (HL1), dimension 66m × 75m (Figure 6.1b). SIS4 is a typical floor plan of an academic building containing a large central structure with elevator shafts, stairwells, and restrooms. Around this center, partitioned by office walls, are classrooms, graduate student offices and a student lounge. HL1 is a large library containing both an open study area and an open area of bookshelves, an area of compact bookshelves, and separated library staff offices. The center of the floor plan consists of stairwells, elevator shafts, and restrooms. The design of the six scenarios described in section 6.1 were considered, in which scenarios 1 through 3 are tailored to the SIS4 environment and scenarios 4 through 6 to the HL1 environment. Different levels of network traffic are considered in each of the scenarios, as described in Table 6.1.

We compare the use of two path loss models: the log distance model and the partition dependent model, for the signal strength calculations in the design process. A detailed description of each model is presented in Section 2.7 and is reviewed below.

1. Log-distance model

The log-distance path loss model for the path loss calculation is written in the equation (6.1).

$$PL(d) = PL(d_0) + 10n\,log_{10}\left(\frac{d}{d_0}\right) + X_\delta \tag{6.1}$$

where d is the distance from the transmitter, $PL(d_0)$ is the path loss at reference distance d_0 , n specifies the path loss behavior for a particular environment, and X_δ represents a normal random variable with standard deviation of δ dB.

In this experiment, n = 3.3 [71] and δ = 3.5 [72] are applied in the path loss calculation using the equation (6.1). To provide 95% signal coverage availability [62], a fading margin of 5.75 dB is applied in the path loss calculation of the design process.

2. Partition dependent model

The partition dependent model [29] considers an explicit number of partitions intervening between the transmitter and the receiver. It assumes the signal attenuation as in the free space (n =2) plus additional path loss increasing with the number of partitions. The path loss model is given by

$$PL(d) = PL(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) + \sum \vartheta_i \varpi_i + X_\delta \qquad (6.2)$$

where ϑ_i is the number of partitions of type i and ϖ_i is the attenuation factor in dB for the partition of type i. The loss values for different types of partitions at 2.4 GHz are reported by Pahlavan and Krishnamurthy [71].

In this experiment, we considered the concrete wall and the elevator shaft to have a loss of 12.4 dB and the floor-to-ceiling walls partitioning classrooms and office spaces to have a loss of 6 dB at 2.4 GHz [71]. In the implementation of the partition dependent path loss model, we count the number of partitions crossed by a straight line connecting the AP and the receiving point. A fading margin of 5.75 is applied to provide 95% signal coverage availability [62], considering δ of 3.5 in both service environments studied here [72].

6.6.2 Numerical Results and Discussions

6.6.2.1 Is the design sensitive to path loss models?

First, the six scenarios considered in this experiment were designed using the log distance path loss model in the signal strength calculation of the design process. Then the resulting network configurations were analyzed using the partition dependent model. The results show that the network configurations designed using the log distance model do not meet all constraints when analyzed using the partition dependent model. From the network design scenarios considered here, it can be concluded that the network design is sensitive to the path loss model used.

6.6.2.2 Effect of Using Different Path Loss Models on the Resulting Network Configurations

For each design scenario, different path loss models were used in the design process. Numerical results show that the use of different path loss model in the network design yield different network configurations as described in the following:

Table 6.9 shows the number of APs employed in the resulting network configurations designed by using different path loss models in the signal strength calculation of the design process. We can see in the low traffic load scenarios of the SIS4 and HL1 service areas (scenarios 1 and 4, respectively), that the partition dependent model yields network configurations employing more APs than configurations designed using the log distance model. In the SIS4 scenario 1, the log distance model yields a network configuration with two APs, while the partition dependent model yield a network with three APs. Figure 6.17 shows network configurations of SIS4 in low network traffic conditions (scenario 1).

Table 6.9 Number of APs used in network configurations

			Number of APs used in the resulting network configurations	
Service area	Scenario	Expected traffic load	Log distance model	Partition dependent model
SIS4	1	low	2	3
	2	medium	3	3
	3	high	4	4
HL1	4	low	5	6
	5	medium	7	7
	6	high	8	8

In the HL1 scenario 4, the log distance model yields a network configuration using five APs, while the partition dependent model yield a network using six APs. Figure 6.20 shows the resulting network configurations of HL1 in low network traffic conditions (scenario 4).

In the network design for medium and high network traffic (scenarios 2, 3, 5 and 6), both path loss models considered in this study yield network configurations employing the same number of APs as shown in Table 6.9. However, the APs' parameters, specifically their locations and power levels, are different, as described below:

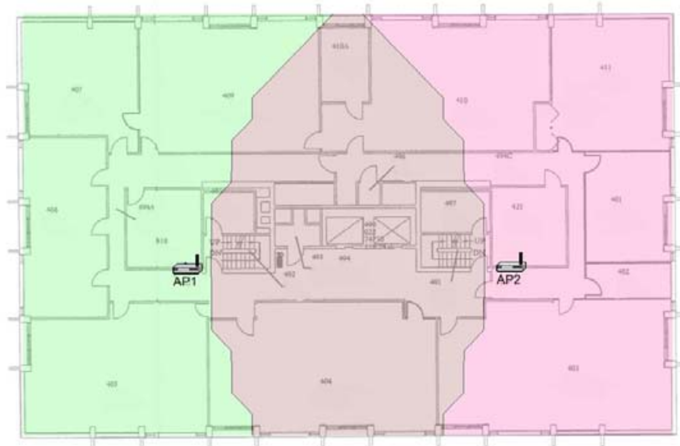
In the SIS4 medium network traffic scenario (scenario 2), both path loss models result in network configurations using three APs. Figure 6.18 shows the resulting network configurations in which the APs' locations are slightly different while the power levels and frequency channels are the same.

In the SIS4 high network traffic scenario (scenario 3), each of the resulting network configurations employs four APs (Figure 6.19). However, they are slightly different from each other with regard to the APs' locations and power levels. Using the partition path loss model, the APs are assigned higher power levels compared to those designed by using the log distance model.

In the HL1 medium and high network traffic scenarios (scenarios 5 and 6), seven and eight APs are used in the resulting network configurations, respectively. Figures 6.21 and 6.22 depict the resulting network configurations in which the APs' locations are different and the power levels are higher in the network configurations designed by using the partition model compared to those designed of the log distance path loss model.

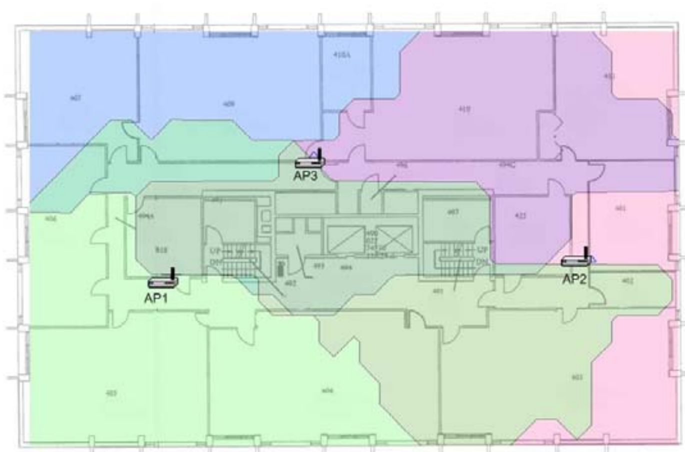
From the results of the numerical experiment described in this section, the following observations and explanations can be made:

- In the low network traffic scenarios (scenarios 1 and 4), the amount of traffic in the network is not a critical issue in the network design, rather the dominant factor is the obstructive characteristics of the service area. The network configurations designed using the path loss model that does not take into account detailed obstructions between the transmitter and the receiver (e.g., the log distance model) utilize fewer APs than the network configurations designed using the path loss model that accounts for detailed obstruction in the service area (e.g., the partition dependent model).
- In the medium and high network traffic scenarios (scenarios 2, 3, 5 and 6), the amount of network traffic is an issue in the network design. Using either the log distance or the partition dependent path loss models in the design process yields WLAN configurations that require the same number of APs. The reason is that APs used to accommodate traffic demand in the service area in turn provide signal coverage to the area where obstructions might exist. Therefore, the resulting network configurations designed by using both path loss models require the same number of APs in the network. However, the APs' parameters (e.g., locations, power levels, and frequency channels) are changed when using the partition dependent path loss model which accounts for detailed obstructions in the service environment.



AP	Location	Frequency channel	Power level
1	(8.0,8.0)	1	6
2	(24.0,8.0)	6	6

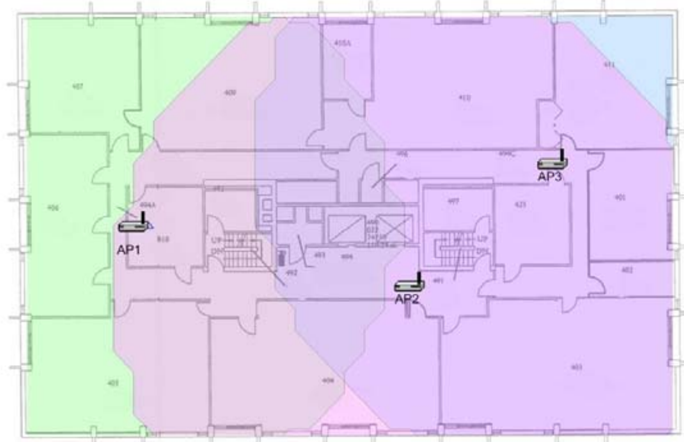
a. WLAN configuration designed by using the log distance path loss model



AP	Location	Frequency channel	Power level
1	(8.0,8.0)	1	6
2	(27.0,10.0)	6	6
3	(14,14)	11	4

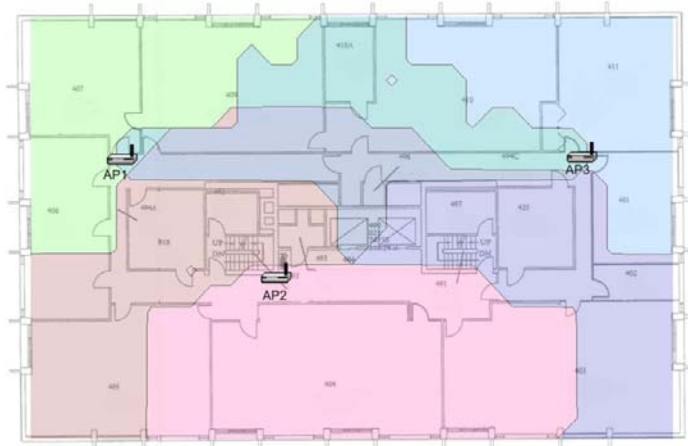
b. WLAN configuration designed by using the partition dependent path loss model

Figure 6.17 Comparison of WLAN configurations for SIS4 with low network traffic (scenario 1)



AP	Location	Frequency channel	Power level
1	(6.0,10.0)	1	6
2	(19.0,7.0)	6	6
3	(26.0,13.0)	11	6

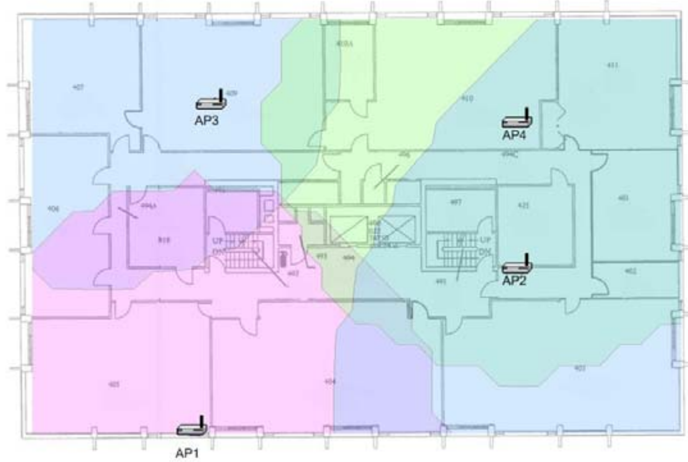
a. WLAN configuration designed by using the log distance path loss model



AP	Location	Frequency channel	Power level
1	(5,13)	1	6
2	(13,8)	6	6
3	(27,13)	11	6

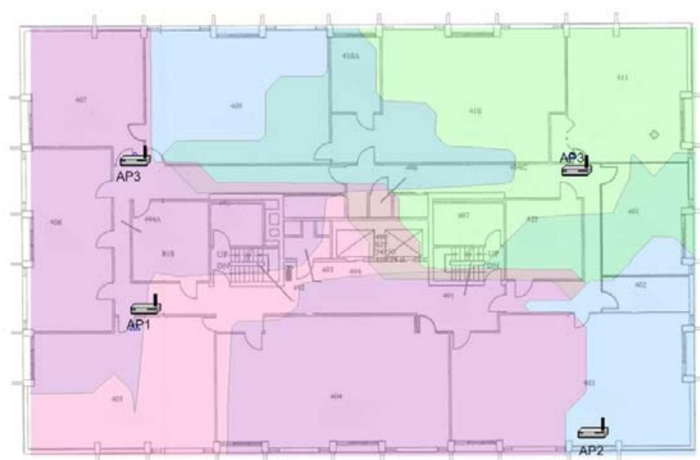
b. WLAN configuration designed by using the partition dependent path loss model

Figure 6.18 Comparison of WLAN configurations for SIS4 with medium network traffic (scenario 2)



AP	Location	Frequency channel	Power level
1	(8.0,0.0)	6	5
2	(24.0,8.0)	1	6
3	(9.0,16.0)	1	5
4	(23.0,15.0)	11	5

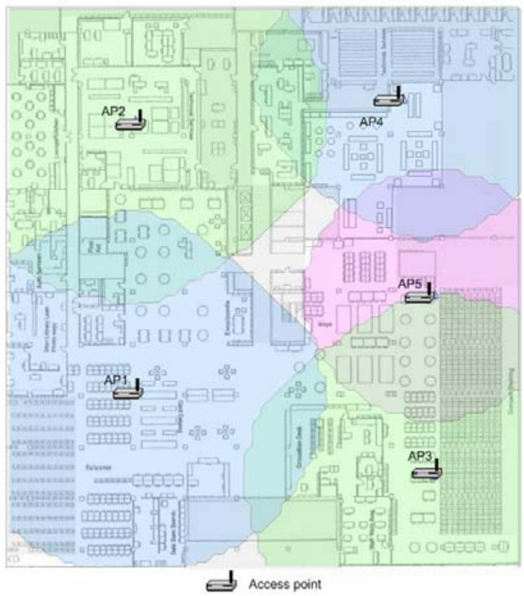
a. WLAN configuration designed by using the log distance path loss model



AP	Location	Frequency channel	Power level
1	(5,6)	6	6
2	(27,1)	1	5
3	(5,14)	1	5
4	(26,14)	11	6

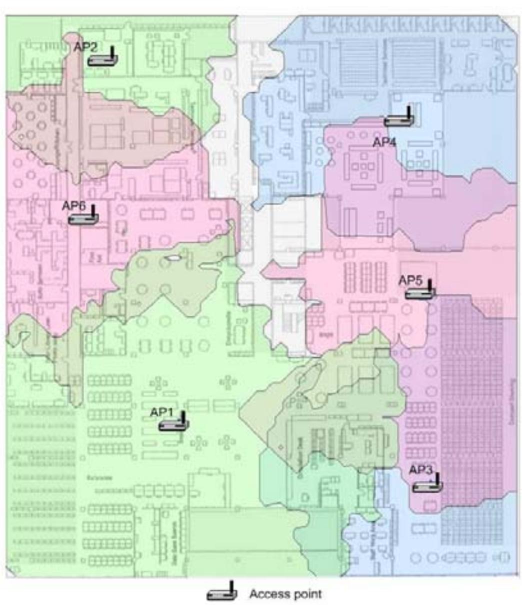
b. WLAN configuration designed by using the partition dependent path loss model

Figure 6.19 Comparison of WLAN configurations for SIS4 with high network traffic (scenario 3)



AP	Location	Frequency channel	Power level
1	(16.0,23.0)	6	6
2	(16.0,59.0)	1	6
3	(54.0,12.0)	1	6
4	(50.0,62.0)	6	5
5	(54.0,36.0)	11	4

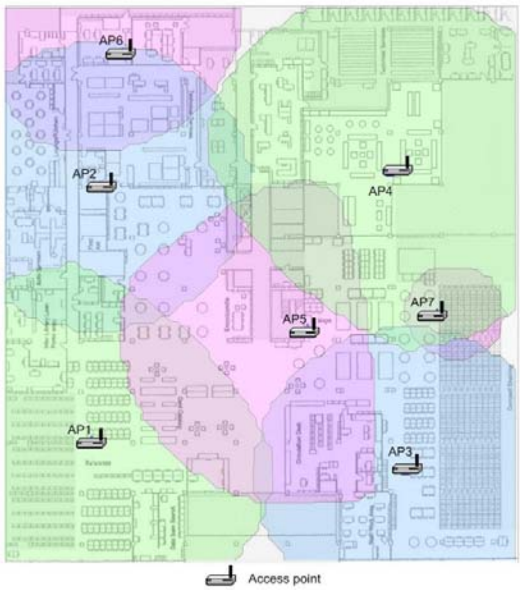
a. WLAN configuration designed by using the log distance path loss model



AP	Location	Frequency channel	Power level
1	(22,20)	1	6
2	(12,68)	1	5
3	(54,12)	6	6
4	(50,60)	6	6
5	(53,37)	11	6
6	(10,47)	11	6

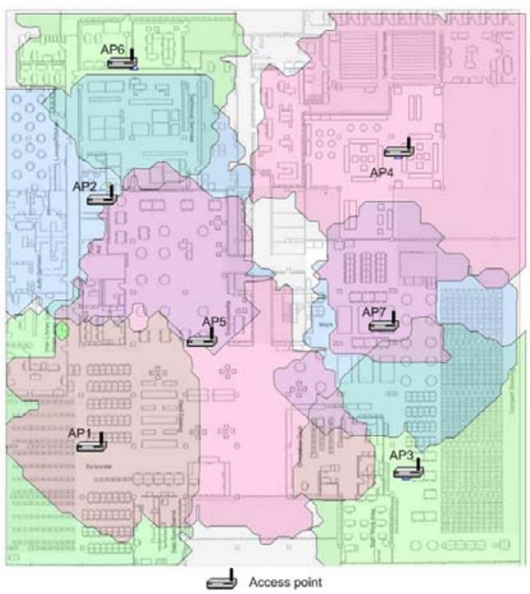
b. WLAN configuration designed by using the partition dependent path loss model

Figure 6.20 Comparison of WLAN configurations for HL1 with low network traffic (scenario 4)



AP	Location	Frequency channel	Power level
1	(11.0,16.0)	1	6
2	(12.0,50.0)	6	5
3	(51.0,12.0)	6	5
4	(49.0,52.0)	1	6
5	(38.0,31.0)	11	6
6	(15.0,68.0)	11	4
7	(55.0,33.0)	11	3

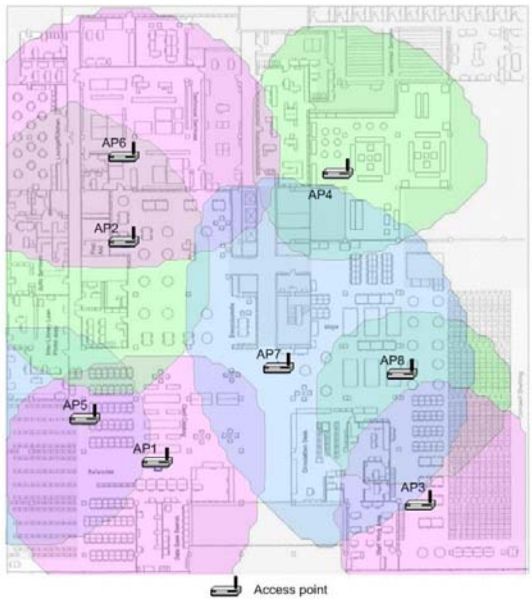
a. WLAN configuration designed by using the log distance path loss model



AP	Location	Frequency channel	Power level
1	(10,16)	1	5
2	(12,49)	6	6
3	(50,12)	1	6
4	(49,55)	11	6
5	(25,30)	11	6
6	(16,67)	1	4
7	(47,32)	6	3

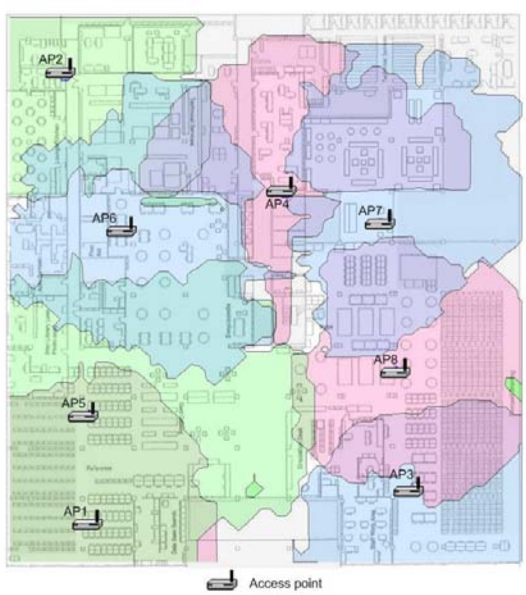
b. WLAN configuration designed by using the partition dependent path loss model

Figure 6.21 Comparison of WLAN configurations for HL1 with medium network traffic (scenario5)



AP	Location	Frequency channel	Power level
1	(18.0,13.0)	11	5
2	(15.0,43.0)	1	5
3	(51.0,10.0)	11	4
4	(40.0,50.0)	1	5
5	(10.0,20.0)	6	5
6	(16.0,52.0)	11	5
7	(35.0,28.0)	6	6
8	(49.0,26.0)	1	4

a. WLAN configuration designed by using the log distance path loss model



AP	Location	Frequency channel	Power level
1	(10,5)	11	5
2	(7,66)	1	6
3	(51,10)	6	6
4	(35,50)	11	6
5	(10,20)	1	6
6	(14,45)	6	6
7	(47,46)	6	6
8	(49,26)	11	4

b. WLAN configuration designed by using the partition dependent path loss model

Figure 6.22 Comparison of WLAN configurations for HL1 with high network traffic (scenario 6)

6.7 Summary

This chapter presents numerical experiments in WLAN design considering various scenarios ranging from the design of a network for a small, single-floor service area to the design of a network for a large, single-floor area to the design of a network for a multiple-floor service area. We also demonstrate the application of the proposed demand-based WLAN design model to an indoor service area being extended to include an outdoor courtyard.

Additionally, a comparative performance evaluation was conducted to compare the performance of WLANs designed by the demand-based network design approach against those designed by the coverage-based design approach. The results demonstrate the superiority of networks designed using the demand-based approach over those designed using the coverage-based approach.

Finally, extensive numerical experiments were conducted to study the effects of shadow fading and the path loss models used in the network design. When using different values for the shadow fading in network design, the resulting network configurations are slightly different from each other. As the shadow fading increases, the power levels of some APs are increased to compensate for greater signal reduction due to signal fluctuation, and the locations of some APs are adjusted to prevent interference between co-channel APs.

The numerical experiments conducted to study the effects of path loss models show that network design is sensitive to the model used in the calculation of the design process. Different path loss models used in the design process yield networks with different configurations. In a service area with low traffic demand, the path loss model considering detailed obstructions between APs and the receiving points yields network configurations requiring more APs than those designed using the distance dependent path loss model. In a service area with high traffic demand, both the partition dependent and the distance dependent path loss models yield network configurations using the same number of APs. However, the APs' parameters are different.

7.0 CONCLUSIONS, CONTRIBUTIONS, LIMITATIONS, AND FUTURE WORK

7.1 Conclusions and Contributions

In this dissertation, we have considered the problem of Wireless Local Area Networks (WLAN) design. As WLANs become more widespread and the price of notebook and handheld computers drops, the number of wireless users and the expected network traffic become issues that need to be considered in the network design process. If these issues are not considered, the resulting networks may not be capable of accommodating their traffic. Existing WLAN design methods limit their focuses to the provision of signal radio coverage in the service area. These are insufficient and inefficient methods for designing networks for current and the emerging WLAN environments in which traffic demand for data communication is high. To address the drawbacks of the existing WLAN design methods, we develop in this dissertation a demand-based WLAN design methodology that accounts for not only the radio signal coverage requirements but also the data rate capacity requirements. Specifically, the major contributions of this dissertation include:

- *Development of a demand-based WLAN design model formulated as a Constraint Satisfaction Problem (CSP).* The design model's mathematical formulation considers both signal radio coverage requirements and data rate capacity requirements. Network usage characteristics of WLANs are accounted for by incorporating the correlation between users' network usage behavior and their locations into the CSP formulation. Additionally, CSMA/CA capacity analytical model is incorporated into the network design to ensure that the average required data rate is available to prospective wireless users in the service area.

- *Development of a heuristic solution technique that can efficiently solve the demand-based WLAN design problem.* The heuristic relies upon simple constraint violation information and can efficiently guide the search to converge to a feasible solution in a reasonable amount of time. Components of the heuristic solution technique framework that this research has developed include: an Area Covering Heuristic (ACH), a Demand Clustering Heuristics (DCH), Frequency Channel Assignment (FCA) based on the simulated annealing algorithm, and the Constraint Violation Reduction (CVR) heuristic based on a tabu search algorithm.
- *The framework of the developed WLAN design methodology is flexible*. Its flexibility due to a modular structure allows for different antenna patterns, path loss models and/or 802.11 capacity analytical models to be used without changing the overall framework of the network design model and the solution technique.
- *The developed demand-based WLAN design methodology is capable of designing various network scenarios.* This methodology is useful for a variety of environments, from those with small, single floor service areas to those that are complex with multiple-floor service areas. The capability of this design methodology is fourfold, as it can determine not only the number of APs required, but also their locations, power levels and frequency channels. Additionally, it can

be applied to the design of service environments that include courtyard areas around the service building.

7.2 Limitations and Future Work

The WLAN design formulation developed in this dissertation incorporates several assumptions which result in limitations in the network design model. Addressing each of these assumptions and their effects is left for future work. The current network design model accounts for only a downlink design (i.e. signal transmitted from APs toward wireless terminals). Interference is calculated based on signal levels from adjacent APs operating on the same or overlapping frequency channels. Interfering signals generated by mobile nodes are not considered. The assumption, based on network trace studies [40], was that most of traffic in a WLAN is in the downlink direction (i.e. data transmitted from APs toward wireless users). Thus, the majority of the interference energy is also from the downlink signals transmitted from APs. In real environment, there might already be APs and occur interference sources in the desired service area, however the current network design model assumes a "greenfield" design. Additionally, the model assumes that wireless users communicate with one AP (i.e. the network design model associates each wireless user to one AP). Therefore traffic load to may be assigned to a specific AP, so that an average obtainable user data rate can be estimated. However, wireless users may be able to communicate with and use the capacity from more than one AP. Finally, the model assumed a constant and uniform data rate within the BSA. However, due to data rate fallback protocols, the data rate may change as signal strength and SIR levels change.

Directions for future work include:

- Incorporating both uplink and downlink signals in the network design formulation.
- Modifying the WLAN design methodology so that it can redesign existing WLANs to accommodate higher expected traffic demand and/or to expand coverage of the present service area.
- Incorporating multiple data rates within the BSA by applying multiple SIR levels in the network design formulation.
- Allowing mobile nodes to multi-home (i.e. associating each wireless user (each demand node) to multiple APs) in order to provide survivability in the WLAN design. If sufficient capacity is designed into the WLAN configuration then when one AP fails, another AP can support the traffic load.

APPENDIX A

List of abbreviations used in the dissertation

ACH	=	Area Covering Heuristic
AP	=	Access Point
BSA	=	Basic Service Area
BSS	=	Basic Service Set
CS	=	Coverage Square
CSMA/CA	=	Carrier Send Multiple Access/ Collision Avoidance
CSP	=	Constraint Satisfaction Problem
CVR	=	Constraint Violation Reduction
DCH	=	Demand Clustering Heuristic
ESS	=	Extended Service Set
FCA	=	Frequency Channel Assignment
ICH	=	Initial Covering Heuristic
MAC	=	Medium Access Control
SA	=	Simulated Annealing
STP	=	Signal Test Point
SIR	=	Signal to Interference Ratio
TS	=	Tabu Search
UA	=	Unit Area
WLAN	=	Wireless Local Area Network

APPENDIX B

List of notations used in the dissertation

(x, y, z)	=	Coordinate in three-dimensional space, where (x, y) represents a location on floor z
A	=	$\{ap_1, ap_2, \ldots, ap_n\}$ denote a set of APs used in the service area, where n is the total number of APs required.
U	=	$\{d_1^t, d_2^t, \ldots, d_m^t\}$ denote a set of demand nodes, given by their position inside the service area, where m is the total number of demand nodes in the set U. Index t indicates the type of sub-area where demand node i is located: $1 \le i \le m$ and $t \in T$ where T is the set of sub-area types.
$\overline{U}$	=	Set of demand nodes that are not associated to any AP
G	=	$\{g_1, g_2, \ldots, g_c\}$ denote a set of STPs representing locations where the received signal strength and the SIR level will be tested, and c is the total number of STPs in the service area. Each STP g_h refers to a coordinate in three-dimensional space (x_h, y_h, z_h), where $1 \le h \le c$ and z_h is the floor where g_h is located.
T	=	Set of sub-area types. Here we define three types of sub-areas: $T = \{1,2,3\}$ where 1 denotes private sub-areas, 2 denotes public sub-areas for unscheduled activities, and 3 denotes public sub-areas for schedule-based activities.
n	=	Number of APs used in the network

p_j = power level of $ap_j \in$ A

f_j = frequency channel of $ap_j \in$ A

d_{ij}^{t} = User association binary variable that equals 1 if demand node $i \in U$ associates to $ap_j \in A$; 0 otherwise. Index $t \in T$ represents the type of sub-area where demand node i is located.

g_{hj} = Signal availability binary variable that equals 1 if STP $h \in G$ can receive a signal from $ap_j \in A$; 0 otherwise.

(x_j,y_j,z_j) = location of $ap_j \in A$, (x_j, y_j) represents the coordinates on floor z_j

D_n = Integer number

D_p = Set of candidate power levels for variable $p_j = \{P_1, P_2,\ldots, P_{max}\}$

D_f = Set of candidate frequency channels for variable $f_j = \{F_1, F_2, \ldots, F_k\}$

D_d = {0, 1} the domain of binary variable d_{ij}^{t}

D_g = {0, 1} the domain of binary variable g_{hj}

$D_{(x,y,z)}$ = The domain of variable (x_j, y_j, z_j) where $x_{min} < x_j < x_{max}$, $y_{min} < y_j < y_{max}$, and $z_j \in FLOOR=\{1^{st}, 2^{nd}, \ldots, r^{th}\}$

α_t = User activity level defines the percentage of wireless users in sub-area type t that are engaged in a data transfer activities (i.e., participating in channel contention and sharing AP capacity)

R_t = Average user data rate requirement in the sub-area type t

$P_{Rthreshold}$ = Received sensitivity threshold

$SIR_{threshold}$ = Signal to interference ratio threshold

η_{jk} = Adjacent channel interference, specifying the percentage of interfering power that the signal operating at f_k dissipates to the signal operating at f_j

C_j = Data rate capacity of the ap_j

$P_{R_{ij}}$ = Received signal strength that user i receives from ap_j

$Intf_{ij}$ = Interference level at user i, when associating to ap_j

r_i^t = Average data rate that user i who situates in sub-area type t can obtain

U_j^t = Set of all prospective wireless users located in sub-area type t that are within the radio signal coverage of ap_j and associate to ap_j

$$\left|U_j^t\right| = \sum_{i \in U} \sum_{j \in A} d_{ij}^t$$

m_j^t = Set of active wireless users located in sub-area type t that are active in data transfer activities and sharing capacity of ap_j. The average number of active users in the set m_j^t is the fraction of users in the set U_j^t that are expected to be active and can be estimated by applying the parameter α_t (user activity level) (i.e., $\left|m_j^t\right| = \alpha_j \left|U_j^t\right|$)

m_j = Set of all active wireless users that are sharing capacity of ap_j

= Union of all active users from all sub-areas that can communicate to ap_j

= $\bigcup_{\forall t \in T} m_j^t$

Λ = Set of all active wireless users in the network = $\bigcup_{\forall j \in A} m_j$

pkt_t = Average packet length (bits) of data traffic from the user in sub-area type t

τ_i^t = Average overall transmission time (μsec) that it takes to successfully transmit a packet from user i located in sub-area type t

$\tau_{overhead}$ = Transmission time according to the CSMA/CA protocol overhead (μsec)

$\tau_{compete}$ = Estimated time spent in the contention period when a group of m_j active users are contending for the data channel at ap_j (μsec)

τ_{pkt} = Packet transmission time (μsec)

$P_c(m_j)$ = proportion of collisions experienced by each packet successfully acknowledged at the MAC layer when a group of m_j active users are contending for the data channel at ap_j ($0 \le P_c(m_j) < 1$)

$DIFS$ = Distributed Interframe Space

$PLCP_{preamble}$ = Preamble part of the Physical Layer Convergence Protocol

$PLCP_{header}$ = Header part of the Physical Layer Convergence Protocol

$SIFS$ = Short Interframe Space

ACK = Acknowledgement

$SLOT$ = Time slot

$CW_{\min}$ = Contention window

CRC = Cyclic Redundancy Check

MAC_{header} = Average medium access layer header

CoG = Center of gravity

γ = Radius of AP coverage area computed using a path loss model and assuming an isotropic antenna.

(x_{CoG}, y_{CoG}) = Coordinate of CoG

Ψ = Graph (V, E) represent a WLAN in which $V = \{v_1, v_2, \ldots, v_n\}$ is the set of vertices representing APs in the set $A = \{ap_1, ap_2, \ldots, ap_n\}$ and E = $\{(v_i,v_j)/\ v_i,v_j \in V)$ is the set of edges representing channel separation constraints between all pairs of APs.

B = $(b_{ij})^{n' \times n'}$ channel separation matrix where n' is the number of vertices of the graph, element b_{ij} with $i \neq j$ representing the minimum channel separation required between vertices v_i and v_j such that the network can avoid interference caused by other APs in the vicinity.

$distance_{max}$ = Channel separation between each pair of consecutive non-overlapping channels

CT = Control temperature in the simulated annealing algorithm

Θ = Cost function of the WLAN FCA problem

$n_{violated}$ = The number of violations on the channel separation requirement in the WLAN FCA problem.

Max_{iter} = Maximum number of iterations allowed at each temperature state in the simulated annealing process of the WLAN FCA problem

P_{accept} = Probability of accepting new frequency assignment in the simulated annealing process of the WLAN FCA problem

φ = the parameter that specifies how the CT is reduced and how the number of iterations at each temperature step is increased as CT is reduced

Max_noimprov = the number of consecutive iterations that allow the search to proceed without improvement

CT_{max} = the starting temperature

CT_{min} = the finishing temperature

S = $\{s_1, s_2, s_3, \ldots, s_m\}$ a set of neighbor solutions generated based on the current solution s.

s^{**} = The best_so_far solutions

s^{*} = The best solution in **S**

β_1 = measures the violations of the radio signal coverage criterion

β_2 = measures violations of the data rate requirement criterion

w_1 = weight factor representing relative importance of radio signal coverage criteria

w_2 = weight factor representing relative importance of traffic demand coverage criteria

w_g = weight factor representing relative importance of grid g

w_i = weight factor representing relative importance of user i

UA = Unit Area in the service area

n_{UA} = Total number of UAs managed in the service area

G_i^{pr} = Set of STPs in UA_i where the received signal strength is below the threshold, for i = 1,…, n_{UA}

G_i^{SIR} = Set of STPs in UA_i where the SIR level is below the threshold, for i = 1,…, n_{UA}

A_i^{ov} = Set of APs in UA_i that cannot provide the required average data rate to associated wireless users for i = 1,…, n_{UA}

CoG_i^{pr} = Center of gravity of STPs in G_i^{pr}

l_i^{pr} = The lowest received signal strength among STPs in G_i^{pr}

Loc_i^{pr} = Location of the STP with the lowest signal strength

$\left\|G_i^{SIR}\right\|$	=	Number of STPs in G_i^{SIR}
avg_i^{SIR}	=	Average SIR level of all STPs in G_i^{SIR}
CoG_i^{SIR}	=	Center of gravity of STPs in G_i^{pr}
$CoG_i^{a^{ov}}$	=	Center of gravity of the locations of wireless users assigned to $a^{ov} \in A_i^{ov}$ where a^{ov} denotes an overload AP
$A_{potential}^{ssv}$	=	Set of potential APs selected with respect to the signal strength
	=	violation (*ssv*)
$A_{potential}^{SIRv}$	=	Set of potential APs selected with respect to the SIR violation (*SIRv*)
$A_{potential}^{drv}$	=	Set of potential APs selected with respect to the average data rate violation (*drv*)
$A_{potential}$	=	Set of all potential APs = $A_{potential}^{ssv} \cup A_{potential}^{SIRv} \cup A_{potential}^{drv}$
$\Omega_{g.m.}^{ssv}$	=	Set of neighbor solutions regarding the signal strength violation derived by the greedy moves (g.m.)
$\Omega_{c.m.}^{ssv}$	=	Set of neighbor solutions regarding the signal strength violation derived by conservative moves (c.m.)
$\Omega_{g.m.}^{SIRv}$	=	Set of neighbor solutions regarding the SIR violation derived by greedy moves
$\Omega_{c.m.}^{SIRv}$	=	Set of neighbor solutions regarding the SIR violation derived by conservative moves
$\Omega_{g.m.}^{drv}$	=	Set of neighbor solutions regarding the average data rate violation derived by greedy moves

$\Omega^{drv}_{c.m.}$ = Set of neighbor solutions regarding the average data rate violation derived by conservative moves

G_i^{SIR} = Set of STPs in UA_i where the SIR level is below the $SIR_{threshold}$, for $i = 1,\ldots, n_{UA}$

G_i^{SIR} is subdivided into σ subsets where

$G_{i,1}^{SIR} \cup G_{i,2}^{SIR} \cup \ldots \cup G_{i,k}^{SIR} \cup \ldots \cup G_{i,\sigma}^{SIR} = G_i^{SIR}$ and

$G_{i,1}^{SIR} \cap G_{i,2}^{SIR} \cap \ldots \cap G_{i,k}^{SIR} \cap \ldots \cap G_{i,\sigma}^{SIR} = \varnothing$

σ = Total number of APs that provide the highest SIR level to STPs in G_i^{SIR}

$a_{i,k}^{h}$ = AP that provides the highest SIR level to STPs in $G_{i,k}^{SIR}$, for $k = 1, \ldots, \sigma$

$a_{i,k}^{INTF}$ = AP that is closest to $a_{i,k}^{h}$ and operates at the interfering frequency channel with respect to the frequency channel assigned to the $a_{i,k}^{h}$. Note that in the case of 802.11b, the interfering frequency channel is the channel where the channel distance to the frequency channel of the $a_{i,k}^{h}$ is less than five. In the case of the 802.11a, the AP whose frequency channel is the same as that of $a_{i,k}^{h}$ is the interfering AP.

$A_{potential}^{SIR,h}$ = Set of potential APs that could improve the received signal strength so that the SIR level can be improved

$A_{potential}^{SIR,INTF}$ = Set of potential APs that could reduce the interference signal so that the SIR level can be improved

$(\Delta x_j, \Delta y_j)$ = step size to move the $a_j \in A_{potential}^{SIR,INTF}$

$(x\_CoG_{i,k}^{SIR}, y\_CoG_{i,k}^{SIR})$ = *(x, y)* coordinate of the center of gravity $CoG_{i,k}^{SIR}$

$\oplus$ = Direction to move $a_j \in A_{potential}^{SIR,INTF}$. (It indicates whether to move away from or toward $CoG_{i,k}^{SIR}$)

UA size = The size of the unit area that is used to partition the service area in order to manage the constraint violation information

λ = Step size to move AP

Max_{iter_CVR} = The number of consecutive iterations without improvement on the constraint violation reduction

$TabuTenure_{di}$ = The tabu tenure of an AP's move direction attribute (the number of iterations during which the AP cannot return back to the direction from which it is moved)

$TabuTenure_p$ = The tabu tenure of an AP's power level attribute (the number of iterations during which the AP's power level cannot be reduced (increased) if it has just been increased (reduced))

Elite_list_size = The number of good solutions recorded in the list

ϑ_i = The number of partitions of type *i*

ϖ_i = The attenuation factor in dB for the partition of type *i*

BIBLIOGRAPHY

[1] In-Stat/MDR, "It's cheap and it works: Wi-Fi brings wireless networking to the masses," in *Technical report*, 2002.

[2] " World wireless LAN markets," Frost & Sullivan, Technical report 2001.

[3] "IEEE 802.11 news," in *http://www.palowireless.com/i802_11/news.asp*, 2003.

[4] A. Sikora, *Wireless personal and local area networks*: John wiley & sons Ltd, England, 2003.

[5] FCC, "Federal Communications Commission http://www.fcc.gov/."

[6] "Information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," in *ISO/IEC 8802-11; ANSI/IEEE Std 802.11, 1999 edn*, 1999.

[7] "Supplement to IEEE standard for information technology telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: high-speed physical layer in the 5 GHz band," in *IEEE Std 802.11a-1999*, 1999.

[8] "Supplement To IEEE Standard For Information Technology- Telecommunications And Information Exchange Between Systems- Local And

Metropolitan Area Networks- Specific Requirements- Part 11: Wireless LAN Medium Access Control (MAC) And Physical Layer (PHY) Specifications: Higher-speed Physical Layer Extension In The 2.4 GHz Band," in *IEEE Std 802.11b-1999*, 2000, pp. i-90.

[9] "IEEE standard for information technology- telecommunications and information exchange between systems- local and metropolitan area networks- specific requirements Part II: wireless LAN medium access control (MAC) and physical layer (PHY) specifications," in *IEEE Std 802.11g-2003 (Amendment to IEEE Std 802.11, 1999 Edn. (Reaff 2003) as amended by IEEE Stds 802.11a-1999, 802.11b-1999, 802.11b-1999/Cor 1-2001, and 802.11d-2001)*, 2003, pp. i-67.

[10] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Transactions on networking*, vol. 8, pp. 785-799, 2000.

[11] G. Anastasi and L. Lenzini, "QoS provided by the IEEE 802.11 wireless LAN to advanced data applications: a simulation analysis," *Wireless Networks*, vol. 6, pp. 99-108, 2000.

[12] T. S. Rappaport, "Mobile radio propagation: Large-scale path loss," in *Wireless communication: Principles & Practice*, 2nd ed, 2002, pp. 105-177.

[13] J. Kabara, P. Krishnamurthy, and D. Tipper, "Capacity based network planning for wireless data networks," *Proc.IST Mobile Communications Summit*.

[14] A. Hills, "Large-scale wireless LAN design," *IEEE Communication Magazine*, vol. 39, pp. 98-104, 2001.

[15] M. Unbehaun, "Scalability of wireless LAN systems in the unlicensed 17 GHz frequency band," *GLOBECOM'01, IEEE*, vol. 6, pp. 3603, 2001.

[16] "Guidelines for IEEE 802.11 cell planning," Lucent Technologies, Feb 1999.

[17] B. J. Bennington and C. R. Bartel, "Wireless Andrew: Building a high speed, campus-wide wireless data network," *Mobile networks and applications*, vol. 6, pp. 9-22, 2001.

[18] M. Unbehaun and J. Zander, "Infrastructure density and frequency reuse for user-deployed wireless LAN systems at 17 GHz in an office environment," *ICC, Helsinki, Finland*, pp. 2535-2539, 2001.

[19] "Nokia Site survey tool," in *www.nokia.com/3g/pdf/stt.pdf*: Nokia Networks, 2001.

[20] "InFielder," in *www.wirelessvalley.com/Products/InFielder/InFielder.asp*: Wireless Valley communications, Inc., 2001.

[21] "ProMan," in *www.awe-communications.com/ProMan.html*: AWE Communications, 2001.

[22] "Predictor," in *www.wirelessvalley.com/Products/Predictor/Predictor.asp*: Wireless Valley communications, Inc., 2001.

[23] R. P. Torres, L. Valle, M. Domingo, and S. Loredo, "CINDOOR: An engineering tool for planning and design of wireless systems in enclosed spaces," *IEEE Antennas and Propagation Magazine*, vol. 41, pp. 11-22, 1999.

[24] M. D. Adickes, R. E. Billo, B. A. Norman, S. Banerjee, B. O. NNaji, and J. Rajgopal, "Optimization of indoor wireless communication network layouts,"

Technical Report, 99-5, Dept.of Industrial Engineering, University of Pittsburgh, 1999, 1999.

[25] W. Tomasi, "Digitial communications," in *Electronic communications systems*, 3 ed, 1998, pp. 431-481.

[26] H. R. Anderson and J. P. McGeehan, "Optimizing microcell base station locations using simulated annealing techniques," *Vehicular Technology Conference, IEEE*, vol. 2, pp. 858-862, 1994.

[27] V. Erceg, "Urban/suburban out-of-sight propagation modeling," *IEEE Communication Magazine*, pp. 56-61, 1992.

[28] H. D. Sherali, C. M. Pendyala, and T. S. Rappaport, "Optimal location of transmitters for micro-cellular radio communication system design," *IEEE Journal on selected areas in communications*, vol. 14, pp. 662-673, 1996.

[29] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communicatons in multi-floored buildings," *IEEE Transactions on Antennas Propagation*, 1992.

[30] D. Stamatelos and A. Ephremides, "Spectral efficiency and optimal base placement for indoor wireless networks," *IEEE Journal on selected areas in communications*, vol. 14, pp. 651-661, 1996.

[31] R. C. Rodrigues, G. R. Mateus, and A. A. F. Loureiro, "On the design and capacity planning of a wireless local area network," *IEEE Conference on Network Operations and Management Symposium*, pp. 335-348, 2000.

[32] M. Unbehaun, "Coverage planning for indoor wireless LAN systems," http://www.s3.kth.se/~matthias/Public/Stanford/MatthiasUnbehaun.Abstract.pdf., 2001.

[33] A. Gamst, E. G. Zinn, R. Beck, and R. Simon, "Cellular radio network planning," *IEEE Aerospace and Electronic Systems*, vol. 1, pp. 8-11, 1986.

[34] F. Perez-Fontan and J. M. Hernando-Rabanos, "Educational cellular radio network planning software tool," *IEEE Transactions on education*, vol. 41, pp. 203-215, 1998.

[35] K. Tutschku, "Demand-based radio network planning of cellular mobile communication systems," *INFOCOM 98*, vol. 3, pp. 1054-1061, 1998.

[36] K. Tutschku and P. Tran-Gia, "Spatial traffic estimation and characterization for mobile communication network design," *IEEE Journal on selected areas in communications*, vol. 16, pp. 804-811, 1998.

[37] P. Tran-Gia, K. Leibnitz, and K. Tutschku, "Teletraffic issues in mobile communication network planning," *Telecommunication Systems*, vol. 15, pp. 3-20, 2000.

[38] Z. Drezner, "Chapter 14," *Facility location: A survey of applications and methods*, 1995.

[39] N. Weicker, G. Szabo, K. Weicker, and P. Widmayer, "Evolutionary multiobjective optimization for base station transmitter placement with frequency assignment," *IEEE Transaction on Evolutionary Computation*, vol. 7, pp. 189-203, 2003.

[40] D. Tang and M. Baker, "Analysis of a local-area wireless network," presented at the sixth annual Int. Conf. on Mobile Computing and Networking (MobiCom), August 2000.

[41] D. Kotz and K. Essien, "Characterizing usage of a campus-wide wireless network," Department of Computer Science, Dartmount College Technical Report TR2002-423, March 2002.

[42] D. Kotz and K. Essien, "Analysis of a campus-wide wireless network," presented at Proc. of the eigth annual Int. Conf. on Mobile Computing and Networking (MobiCom), September 2002.

[43] R. Hutchins and E. W. Zegura, "Measurements from a campus wireless network," presented at IEEE International Conference on Communications (ICC 2002), May 2002.

[44] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan, "Characterizing user behavior and network performance in a public wireless LAN," presented at ACM SIGMETRICS'02, June 2002.

[45] M. Balazinska and P. Castro, "Characterizing mobility and network usage in a corporate wireless local-area network," presented at International Conference on Mobile Systems, Applications, and Services (MobiSys'03), San Francisco, CA, USA, May 2003.

[46] D. Chalmers and M. Sloman, "A survey of quality of service in mobile computing environments," *IEEE Communications Surveys,* http://www.comsoc.org/pubs/surveys, 1999.

[47] J. Weinmiller, H. Woesner, and A. Wolisz, "Analyzing and improving the IEEE 802.11 MAC protocol for wireless LANs," *Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pp. 200-206, 1996.

[48] A. Ayyagari, Y. Bernet, and T. Moore, "IEEE 802.11 Quality of Service," *White paper*, 2000.

[49] C.-F. Hsu, Y.-H. Huang, and T.-D. Chiueh, "Design of an OFDM receiver for high-speed wireless LAN," *IEEE international symposium on circuits and systems*, vol. 4, pp. 558-561, 2001.

[50] T.-D. Chiueh and S.-M. Li, "Trellis-coded complementary code keying for high-rate wireless LAN systems," *IEEE Communications Letters*, vol. 5, pp. 191-193, 2001.

[51] W. Carney, "IEEE 802.11g New draft standard clarifies future of wireless LAN," *White paper*, 2002.

[52] E. Tsang, *Foundations of Constraint Satisfaction*. London: Academic Press, 1993.

[53] K. Nonobe and T. Ibaraki, "A tabu search approach to the constraint satisfaction problem as a general problem solver," *European Journal of Operation Research*, vol. 106, pp. 599-623, 1997.

[54] H. Kanoh, M. Matsumoto, and S. Nishihara, "Genetic algorithms for constraint satisfaction problems," *IEEE International Conference on Intelligent Systems for the 21st Century*, vol. 1, pp. 626-631, 1995.

[55] A. J. Motley and J. M. P. Keenan, "Personal communication radio coverage in buildings at 900 MHz and 1700 MHz," *Electronics Letters*, vol. 24, pp. 763-764, 1998.

[56] J. B. Andersen, T. S. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Communication Magazine*, pp. 42-49, 1995.

[57] N. Bhat, "An antenna specific site modeling tool for interactive computation of coverage regions for indoor wireless communications," in *Electrical Engineering*. Blacksburg: Virginia Polytechnic Institute and State University, 1998.

[58] W. L. Stutzman and G. A. Thiele, *Antenna theory and design*, 2nd ed: John Wiley & Sons, Inc., 1998.

[59] C. A. Balanis, *Antenna theory*: John Wiley & Sons, Inc., 1997.

[60] "802.11 Antenna," http://www.argospress.com/Resources/antennas/80211antenn.htm.

[61] R. S. Elliott, *Antenna theory and design*. Englewood Cliffs, N.J.: Prentice-Hall, 1981.

[62] W. C. Y. Lee, *Mobile communications engineering : theory and applications*, 2nd ed. New York: McGraw-Hill, 1998.

[63] L. Z. Ribeiro and L. A. Dasilva, "A framework for the dimensioning of broadband mobile networks supporting wireless internet services," *IEEE Wireless Communications*, pp. 6-13, June 2002.

[64] P. Oppenheimer, *Top-down network design*: Macmillan technical publishing, 1999.

[65] OPNET, "Understanding WLAN Model Internals and Interfaces," presented at OPNETWORK 2003, Washington, D.C., 2003.

[66] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement," presented at Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002), June 2002.

[67] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance Anomaly of 802.11b," presented at Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), April 2003.

[68] C. Prommak, J. Kabara, D. Tipper, and C. Charnsripinyo, "Next generation wireless LAN system design," presented at MILCOM 2002. Proceedings, 2002.

[69] S. C. Johnson, "Hierarchical Clustering Schemes," *Psychometrika*, vol. 2, 1967.

[70] L. Technologies, "Guidelines for IEEE 802.11 cell planning," *WaveLAN Technical Bullentin 024/B*, 1999.

[71] K. Pahlavan and P. Krishnamurthy, "Path-loss modeling and signal coverage," in *Principles of Wireless Networks: A Unified Approach*: Prentice-Hall, Inc., 2002, pp. 46-58.

[72] G. J. M. Janssen, P. A. Stigter, and R. Prasad, "Wideband indoor channel measurements and BER analysis of frequency selective multipath channels at 2.4, 4.75, and 11.5 GHz," *Communications, IEEE Transactions on*, vol. 44, pp. 1272-1288, 1996.

[73] N. Dunkin and S. Allen., "Frequency assignment problems: representations and solutions," Department of Computer Science, Royal Holloway, University of London, Egham, Surrey, UK, TR CSD-TR-97-14, 1997.

[74] S. U. Thiel, S. Hurley, and D. H. Smith, "Frequency assignment algorithms," Dept. of Computer science, University of Wales Cardiff, UK, Ref. RCCM 070, April 1997.

[75] L. E. Hodge, S. Hurley, and D. H. Smith, "Higer-order constraint techniques for the frequency assignment problem," Cardiff University, Cardiff, Wales, U.K. Jan 2002.

[76] K. I. Aardal, S. P. M. V. Hoesel, A. M. C. A. Koster, C. Mannino, and A. Sassano, "Models and solution techiques for frequency assignment problems," Konrad-Zuse-Zentrum fur Informationstechnik, Berlin, Germany ZIB-report 01-40, 2001.

[77] C.-W. Sung and W.-S. Wong, "A graph theoretic approach to the channel assignment problem in cellular systems," presented at IEEE 45th Vehicular Technology Conference, July 1995.

[78] J. Janssen, K. Kilakos, and O. Marcotte, "Fixed preference channel assignment for cellular telephone systems," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 533 -541, March 1999.

[79] S. Hurley, D. H. Smith, and S. U. Thiel, "FASoft: A System for Discrete Channel Frequency Assignment," *Radio Science*, vol. 32, pp. 1921-1939, 1997.

[80] J.-K. Hao, R. Dorne, and P. Galinier, "Tabu Search for Frequency Assignment in Moblie Radio Networks," *Journal of Heuristics*, vol. 4, pp. 47-62, 1998.

[81] D. Costa, "On the use of some known methods for T-colorings of graphs," *Annals of Operations Research*, vol. 41, pp. 343-358, 1993.

[82] D. Beckmann and U. Killat, "A New Strategy for the Application of Genetic Algorithms to the Channel-Assignment Problem," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 1261-1269, 1999.

[83] z. Michalewicz and D. B. Fogel, *How to solve it: Modern Heuristics*: Springer-Verlag Berlin Heidelberg, Germany, 2000.

[84] F. Glover and M. Laguna, *Tabu Search*. Massachusetts: Kluwer Academic Publishers, 1997.

[85] *Webster's II New Riverside Dictionary*: Berkley Books of New York City, 1984.

[86] S. Kirkpatrick, "Optimization by simulated annealing: Quantitative studies," *Journal of Sataistical Physics*, vol. 34, pp. 975-986, 1984.

[87] CISCO, "Channels and power levels," in *Cisco aironet 1200 series access point configuration guide*. San Jose, CA, 2001.

[88] B. Alexander and S. snow, "Preparing for Wireless LANs," in *Packet magazine, Cisco systems*, vol. 14, 2003.

[89] A. V. Pais, K. W. Sowerby, and M. J. Neve, "Outdoor-to-indoor signal correlation and its influence on indoor system performance at 2 GHz," *Electronics Letters*, vol. 39, pp. 236-238, 2003.

[90] A. Chandra, "Propagation of 2000 MHz radio signal into a multi-storeyed building through outdoor-indoor interface," presented at The 14th IEEE 2003 International Symposium on Personal, Indoor and Mobile Communication, 2003.

[91] G. Durgin, T. S. Rappaport, and H. Xu, "Measurements and models for radio path loss and penetration loss in and around homes and trees at 5.85 GHz," *IEEE Transaction on Communications*, vol. 46, pp. 1484-1496, 1998.

[92] R. Gandhi, "Empirical path loss model for outdoor 802.11b wireless links," Indian Institute of Technology, India, 2003.