Capacity Enhancement of Mobile Radio Networks using Intermediate Relays

vorgelegt von

Seble Mengesha Aberra
(M.Sc in Electrical Engineering)

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
– Dr.-Ing. –
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Dr. Holger Boche
Gutachter:	Prof. Dr.-Ing. Adam Wolisz
Gutachter:	Prof. Dr.-Ing Jochen Schiller

Tag der wissenschaftlichen Aussprache: 02.Mai 2006

Berlin 2006

D 83

Zusammenfassung

In derzeitigen infrastrukturbasierten zellularen drahtlosen Netzwerken kommunizieren mobile Teilnehmer, so genannte Terminals, direkt mit ihrem Access-Point, und der eine globale Verbindungen durch die existierende Infrastruktur ermöglicht. Wenn immer mehr Terminals im drahtlosen Netzwerk aktiv werden, wird es vorkommen, dass Terminals gleichzeitig im selben Frequenzband ihre Daten übertragen wollen, was eine Erhöhung der Interferenzen im Gesamtnetz zur Folge hat. Außerdem benutzen die Terminals relativ hohe Sendeleistungen, was zu Interferenzen mit dem restlichen Funkverkehr führt. Zwangsläufig sind die spektrale Effizienz und die maximal erreichbare Kapazität des Netzwerkes durch die interferenzbedingten Störungen begrenzt. In dieser Arbeit wird das Weiterreichen von Daten aus dem Bereich der ad-hoc Netzwerke übernommen und auf Infrastrukturnetze angewendet. Somit werden die Teilnehmer eines drahtlosen Netzes selber als Weiterleiter von Daten eingesetzt und die Gesamtkapazität des Netzwerkes wird erhöht. Das Grundprinzip für die Erhöhung der Kapazität ist, dass mit Hilfe von Weiterleitern die Kommunikationsdistanz zwischen einem entfernten Teilnehmer und seinem Zugangspunkt verringert wird und deshalb höhere Datenraten zum Einsatz kommen können.

Im Mittelpunkt dieser Arbeit steht zunächst die Abschätzung des potentiellen Nutzens des entworfenen drahtlosen Netzwerkes durch getrennte Untersuchung der begrenzenden Faktoren Sendeleistung und Interferenz. Auf Basis dieser ersten Resultate werden dann Methoden zur Erhöhung der Netzwerkkapazität entwickelt: Terminals entscheiden sich, ob sie einen Weiterleiter benötigen um ihre Daten zu übertragen. Dadurch wird ein Routing-Problem gelöst. Ein Weiterleiter wird mit optimaler Sendeleistung und -rate gewählt, nachdem die Entscheidung zur Weiterleitung gefallen ist. Somit wird das Problem der Anpassung der Übertragungsenergie und -rate gelöst.

Drei Wege zur gerechten Verteilung der Resourcen werden vorgeschlagen: das Modell gleicher Slotgrößen, das Modell der gleich bleibenden Datenverkehrsgröße, und nachfrage-orientiertes Schema. Das Problem der Slotzuordnung wird dann gelöst, um die Kapazität weiter zu verbessern. Außerdem werden Methoden vorgestellt, um innerhalb einer Zelle zwei oder mehr Frequenzen neben der primären Frequenz der Zelle zu nutzen. Diese zusätzlichen Frequenzen werden zur Vermittlung genutzt und man führt eine Wiederbenutzung der Frequenzen benachbarten Zellen durch. Daraus resultiert ein möglicher Zuwachs der Netzkapazität von bis zu 60%. Zusätzlich zu diesem Gewinn erreicht man, dass bei steigender Anzahl von Terminals der Durchsatz pro Terminal nicht gegen Null geht, sondern sich pro Knoten ein konstanter Durchsatz erzielen lässt.

Abstract

In infrastructure-based cellular wireless networks, mobile terminals communicate directly with their access point and the access point establishes global connectivity through the infrastructure network. When more terminals are squeezed into the network, there will be terminals transmitting simultaneously in the same frequency band causing interference in the network. In addition, remote terminals communicate with their access point using relatively high transmission power which cause interference to other transmitting terminals. Inevitably, the spectral efficiency and the maximum achievable capacity of the network is limited due to the ensuing interference. In this thesis, the relaying feature of terminals in ad hoc networks is adopted in the infrastructure-based cellular wireless network so that terminals can be used as intermediate relays to improve the capacity of the network. The rationale is that the intermediate terminals reduce the communication distance and allow higher data rates.

The focus of this thesis is to first assess the potential benefit of the proposed network by separately studying the network performance limiting factors–the transmission power and interference– as transmission power limited and interference limited system. Based on these initial results, methods of increasing the network capacity are then developed: Terminals decide if they need an intermediate relay terminal to send their traffic, i.e., a routing problem is solved. A relaying terminal is selected with an optimal transmission power and rate after the relaying decision, i.e, transmission power and rate adaptation problem is solved.

Different ways of fairly sharing the network resources are also suggested: The uniform slot size scheme, when terminals obtain equal share of the total TDMA frame, the uniform traffic size scheme, when terminals are allowed to send a uniform amount of data in slots of varying length and demand oriented scheme when terminals partially share the TDMA frame according to their traffic size and partially obtain an equal time share. The scheduling problem is thence solved to further improve the capacity. Furthermore, methods of using two or more frequencies within a cell, in addition to the cell's primary frequency, are presented. These additional frequencies are used for relaying and are obtained by "recycling" from neighboring cells. As a result it is possible to achieve up to 60% increase in network capacity. And as the number of terminals increase, it is possible to achieve a per node throughput that will not actually go to zero but become a constant.

Acknowledgment

First and foremost I would like to thank Prof. Dr.-Ing Adam Wolisz for his supervision, discussion, suggestion and help in every possible way. Without his advice and reassurance, this work would not have seen the light of day. My sincere thanks also goes to Prof. Dr. rer.nat. Holger Karl for his fruitful discussion and supervision that greatly enriches the thesis. I am thankful to the TKN members for the friendly and exciting research atmosphere and cooperation. I am especially grateful to Daniel Hollos, Martin Kubisch and Andreas Willig. Many other people in different walks of life have supported me along the way, not to mention Dn. Ephrem, Family Ayalneh, Family Tessema, Hani, Jörg, Selamawit, Tsion and Rahel. I am very much thankful to them. I am also thankful to the German Academic Exchange Service (DAAD) for the scholarship. Finally, my heartfelt indebtedness goes to my family and Emnete for their love, support and encouragement throughout my study.

Contents

Zusammenfassung **iii**

Abstract **v**

Acknowledgment **vii**

Acronyms **xxi**

1 Introduction **1**

1.1 Motivation 1
1.2 Problem Statement and Thesis Contribution 2
1.3 Thesis Structure 4

2 Background **7**

2.1 Introduction 7
2.2 Channel Model 7
  2.2.1 Pathloss 8
  2.2.2 Fading 9
2.3 WLAN Technologies 9
  2.3.1 IEEE 802.11 10
  2.3.2 HiperLAN/2 14
2.4 Multi-hop Systems 16
2.5 Medium Access 17
  2.5.1 TDMA 17
  2.5.2 CDMA 18
  2.5.3 OFDM 18
2.6 Frequency Reuse 19
2.7 Transmission Power and Interference Limited Systems 20
  2.7.1 Transmission Power Limited 20
  2.7.2 Interference Limited 20

3 System Model

3.1 Introduction . . . 23
3.2 Channel Model . . . 24
3.3 System Entities . . . 26
  3.3.1 Cell . . . 26
  3.3.2 Access Points . . . 28
  3.3.3 Terminals . . . 28
3.4 System Parameters . . . 29
  3.4.1 Transmission Power . . . 29
  3.4.2 Interference . . . 30
  3.4.3 Signal to Interference and Noise Ratio . . . 30
  3.4.4 Receiver Sensitivity . . . 31
3.5 Communication Protocol . . . 32
  3.5.1 Physical Layer . . . 32
  3.5.2 MAC Structure . . . 32
  3.5.3 Central Organization . . . 32
3.6 Load Model . . . 33
3.7 System Metrics . . . 33

4 Related Work **37**

4.1 Introduction . . . 37
4.2 Information Theory Approach . . . 37
  4.2.1 Ad Hoc Networks . . . 37
  4.2.2 Hybrid Networks . . . 39
  4.2.3 Multi-Hop, Multi-channel and Multi-rate Approaches . . . 40
4.3 Power Control . . . 42
4.4 Fairness and Scheduling Approach . . . 44
4.5 Routing, Interference and System Performance . . . 45
4.6 Summary and Outlook . . . 46

5 Analytical Estimation of Cell Capacity **49**

5.1 Introduction . . . 49
5.2 Transmission Power Limited System . . . 50
  5.2.1 Model . . . 50
  5.2.2 Analysis . . . 52
  5.2.3 Numerical Results . . . 56
  5.2.4 Summary . . . 59
5.3 Interference Limited System . . . 60
  5.3.1 Model . . . 60

    5.3.2 Analysis . . . 61
    5.3.3 Numerical Results . . . 64
    5.3.4 Summary . . . 66
  5.4 Probability Estimation of Finding Relaying Terminals . . . 66
    5.4.1 Model and Assumptions . . . 66
    5.4.2 Numerical Results . . . 68

6 Approach **71**
  6.1 Introduction . . . 71
  6.2 Transmission Power and Rate Adaption . . . 72
  6.3 Routing . . . 73
    6.3.1 Relay Selection . . . 76
    6.3.2 Relaying Modes . . . 79
  6.4 Fairness Requirement . . . 81
  6.5 Scheduling . . . 82
  6.6 Frequency Recycling Pattern . . . 85
  6.7 Example Scenario . . . 85

7 The Simulation Tool **87**
  7.1 Introduction . . . 87
  7.2 Simulator Overview . . . 87
  7.3 Input Block . . . 88
    7.3.1 Topology Generator . . . 88
    7.3.2 Load Model . . . 90
    7.3.3 Simulation Parameters and other Inputs . . . 90
  7.4 Channel Model . . . 91
  7.5 The Physical Layer Block . . . 92
    7.5.1 PER Mapping . . . 92
    7.5.2 Modulation Selection . . . 92
    7.5.3 Transmission Power Calculation . . . 93
  7.6 Simulation Engine . . . 93
  7.7 Evaluation Block . . . 93
  7.8 Summary . . . 95

8 Simulation Results **97**
  8.1 Introduction . . . 97
  8.2 Single Cell Network . . . 97
  8.3 Multi-Cell Network . . . 99

9 Impact of Mobility **123**
9.1 Introduction 123
9.2 Mobility Model 124
9.3 Analysis of Relay Utilization due to Mobility 125
9.3.1 Network Model 126
9.3.2 Probability Derivation for Finding a Mobile Relay Terminal 127
9.3.3 Numerical Results 131
9.4 Simulation Setup 133
9.5 Performance Evaluation 133
9.6 Summary 135

10 Comparison: Optimizing Capacity vs. Energy **137**
10.1 Introduction 137
10.2 Evaluation of Energy 138
10.3 Simulation Results and Discussion 142

11 Conclusion, Open Issues and Future Work **147**

Publications **149**

List of Figures

2.1 A two-ray model for radio propagation over a flat reflecting surface 8
2.2 WLAN Configuration (a) ad hoc networking; (b) infrastructure-based. 10
2.3 IEEE 802.11 protocol architecture 11
2.4 IEEE 802.11 medium access and inter-frame space [100] 12
2.5 IEEE 802.11 contention window and wait time [100] 12
2.6 Hidden-terminal problem 12
2.7 Exposed-terminal problem 12
2.8 IEEE 802.11 hidden-terminal provision for collision free access [100] 13
2.9 HiperLAN/2 protocol architecture 14
2.10 HiperLAN/2 MAC frame structure 15
2.11 Minimum reuse distance [84] 19

3.1 Pathloss with two-ray model 25
3.2 Relationship between maximum cell radius, PER and pathloss coefficient with the lowest transmission rate 27
3.3 PER to modulation and SINR mapping [61] 31
3.4 typical fixed TDMA frame 32

5.1 System scenario 50
5.2 Limited Transmission Power and PER as a function of terminal separation for different combinations of α and NBR 53
5.3 Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -75 dBm 56
5.4 Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -80 dBm 56
5.5 Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -90 dBm 56
5.6 Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -100 dBm 56
5.7 Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -75 dBm 57
5.8 Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -80 dBm 57
5.9 Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -90 dBm 57
5.10 Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -100 dBm 57
5.11 Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -75 dBm 58
5.12 Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -80 dBm 58

5.13 Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -90 dBm 58
5.14 Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -100 dBm 58
5.15 Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -75 dBm 59
5.16 Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -80 dBm 59
5.17 Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -90 dBm 59
5.18 Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -100 dBm 59
5.19 Scenario description 61
5.20 Ratio of total throughput for $\alpha = 2.5$ 65
5.21 Ratio of total throughput for $\alpha = 3.0$ 65
5.22 Ratio of total throughput for $\alpha = 3.8$ 65
5.23 Ratio of total throughput for $\alpha = 4$ 65
5.24 A simple example of intercell relaying. Relaying can be beneficial for communicating over a longer distance or around an obstacle 67
5.25 Relay region (left); Detail (right) 68
5.26 Behavior of $P(N, D)$ for various D as N becomes large (graph shows curves for $N = 4, 8, 16, \ldots, 4196$) 69

6.1 Effective data rate and transmission power with respect to distance for 1 % PER and $\alpha = 3.2$ 73
6.2 (a) Number of hops needed to send traffic from the far terminal to the AP; (b) Communication time shared among the terminals involved when relaying in one-frequency is used 74
6.3 End-to-end throughput as a function of source-destination distance for k number of hops; 1% PER and $\alpha = 3.2$ 75
6.4 End-to-end throughput for a cell radius of d; 1% PER and $\alpha = 3.2$ 75
6.5 slot for direct and relay transmission 77
6.6 An example cell where terminals are identified by a sector 79
6.7 An example frequency recylcing 80
6.8 An example of sectored frequency recycling for 7-frequency reuse pattern 81
6.9 Time-freqency representation of the two fairness schemes, with hypothetic terminals A, B and C; A relaying via B 83
6.10 TDMA frame with (a) uniform slot-size and (b) uniform traffic size 84
6.11 A simple cellular network scenario 86

7.1 Simulator architecture 88
7.2 Cell layout with 7-frequency reuse pattern 90
7.3 Cell layout with 19- frequency reuse pattern 90
7.4 An exponential curve fitting for all physical layer modes 92

8.1 Average throughput for single cell as a function of α for a *uniform slot size* fairness scheme and a target PER of 1% 98
8.2 Average throughput for single cell as a function of α for a *uniform traffic* fairness scheme and a target PER of 1% 99
8.3 Average throughput for single cell as a function of α, for a *uniform slot size* fairness scheme and two-frequency relaying, for target PERs of 1%, 3%, 5% and 10% 100
8.4 Average throughput for single cell as a function of α, for a *uniform traffic* fairness scheme and two-frequency relaying, for target PERs of 1%, 3%, 5% and 10% 100
8.5 Average throughput for single cell as a function of terminals in cell for a uniform slot size fairness scheme, pathloss coefficient $\alpha = 3.2$ and target PER of 1% 101
8.6 Average throughput for single cell as a function of terminals in cell for a uniform traffic fairness scheme, pathloss coefficient $\alpha = 3.2$ and target PER of 1% 101
8.7 Average throughput for single cell as a function of terminals in a cell for a uniform slot size fairness scheme, two-frequency relaying, pathloss coefficient $\alpha = 3.2$, and target PERs of 1%, 3%, 5% and 10% 102
8.8 Average throughput for single cell as a function of terminals in a cell for a uniform traffic fairness scheme, two-frequency relaying, pathloss coefficient $\alpha = 3.2$, and target PERs of 1%, 3%, 5% and 10% 102
8.9 multicell configuration 103
8.10 An example of macro-micro frequency assignment 104
8.11 Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 1% PER and 7 frequency reuse pattern and uniform slot-size fairness 105
8.12 Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 5% PER and 7 frequency reuse pattern and uniform slot-size fairness 105
8.13 Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 2-frequency relaying at 1%, 3%, 5% and 10% PERs and 7 frequency reuse pattern and uniform slot-size fairness 106
8.14 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7 frequency reuse pattern and uniform slot-size fairness 106
8.15 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7 frequency reuse pattern and *demand oriented* fairness 107
8.16 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 3 frequency reuse pattern and uniform slot-size fairness 107
8.17 Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 1% PER and 3 frequency reuse pattern and uniform slot-size fairness 108
8.18 Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 5% PER and 3 frequency reuse pattern and uniform slot-size fairness 108

8.19 Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 2-frequency relaying at 1%, 3%, 5% and 10% PERs and 3 and uniform slot-size fairness frequency reuse pattern 109
8.20 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 3 frequency reuse pattern and *demand oriented* fairness 109
8.21 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern and uniform slot-size fairness 110
8.22 Average throughput as a function of varying α, 20 number of terminals, 1% PER, 19 frequency reuse pattern and uniform slot-size fairness 111
8.23 Average throughput as a function of varying α, 20 number of terminals, 5% PER, 19 frequency reuse pattern and uniform slot-size fairness 111
8.24 Average throughput of 2-frequency relaying as a function of varying α, maximum allowed target PER, 20 number of terminals, 19 frequency reuse pattern and uniform slot-size fairness 112
8.25 Packet error rate variation for the 7 modulation rates used in HiperLAN/2 at a distance of 50 m 112
8.26 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern and *uniform traffic-size* fairness 113
8.27 Average throughput as a function of varying α for two-frequency relaying mode, 20 number of terminals, 19 frequency reuse pattern, *uniform traffic-size* fairness and for 1%, 5% and 10% PERs 113
8.28 Average throughput as a function of varying α, 20 number of terminals, 1% PER, 19 frequency reuse pattern and *uniform traffic-size* fairness 114
8.29 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern and *demand oriented* fairness 114
8.30 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7+7 frequency reuse pattern and uniform slot-size fairness 115
8.31 Comparisons of average throughput as a function of terminals per cell for different frequency reuse patterns, 1% PER and $\alpha = 3.2$ for uniform slot-size fairness 116
8.32 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7+7 frequency reuse pattern and *uniform traffic-size* fairness 116
8.33 Average throughput as a function of varying α for two-frequency relaying mode, 20 number of terminals, 7+7 frequency reuse pattern, *uniform traffic-size* fairness and for 1%, 5% and 10% PERs 117
8.34 Average throughput as a function of varying α, 20 number of terminals, 1% PER, 7+7 frequency reuse pattern and *uniform traffic-size* fairness 117
8.35 Average throughput as a function of varying α, 20 number of terminals for two-frequency relaying mode, 3+3 frequency reuse pattern and uniform slot-size fairness 118

8.36 Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$, 3+3 frequency reuse pattern and uniform slot-size fairness 118
8.37 Average throughput of sectored-frequency relaying as a function of varying terminals, maximum allowed target PER, $\alpha = 3.2$ for 3 frequency reuse pattern and uniform slot-size fairness 119
8.38 Average throughput of sectored-frequency relaying as a function of varying terminals, maximum allowed target PER, $\alpha = 3.2$ for 7 frequency reuse pattern and uniform slot-size fairness 120
8.39 Average throughput of sectored-frequency relaying as a function of varying terminals, maximum allowed target PER, $\alpha = 3.2$ for 19 frequency reuse pattern and uniform slot-size fairness 120
8.40 Per-node average throughput as a function of terminals per cell for different frequency reuse patterns and 1% PER, $\alpha = 3.2$ 122

9.1 Traveling pattern of a mobile terminal with a random walk mobility model 125
9.2 Simple single-relay-hop network with relay region 126
9.3 Possible directions that a mobile terminal x can move towards the AP 128
9.4 The mobile terminals in rectangular coordinate 129
9.5 The mobile terminals in polar coordinate 129
9.6 Probability of finding a usable relay mobile terminal as a function of terminals in a cell for $D = 3.5$ 131
9.7 Probability of mobile terminals as a function of cell and transmission radius ratio . . 132
9.8 Average throughput at the AP for 20 terminals per cell 134
9.9 Average throughput at the AP for 20 mobile terminals with varying terminal speed . 135

10.1 Power associated with the different operation modes of a wireless terminal 139
10.2 Power consumption model 139
10.3 Energy consumption per correct bit for a single terminal sending a fixed amount of data to the AP 140
10.4 An example scenario with corresponding modulation and transmission power for a pathloss coefficient $\alpha = 3.2$ 142
10.5 Communication schedule for (a) direct (b) one-frequency and (c) two-frequency . . . 142
10.6 The total energy consumed in a cellular network as a function of the mobile terminals. 143
10.7 Ratio of total energy consumed in direct communication to relaying in a cellular network as a function of the mobile terminals. 144
10.8 Total energy consumed by 20 mobile terminals in a cellular network for different pathloss coefficients 145
10.9 Total energy consumed by 20 mobile terminals in a cellular network for different target PER 145

List of Tables

2.1 Physical layer modes of HiperLAN/2 [61] 14
2.2 Physical layer modes of IEEE 802.11a [84] 14

3.1 Initial model parameters [91] 25
3.2 SNR and receiver sensitivity [29] 31

5.1 Parameters for C/I to PER interpolation 51
5.2 Direct communication schedule (uplink) 52
5.3 Relaying communication schedule (uplink) 54
5.4 Schedule for direct communication 61
5.5 Schedule for relay communication 62

Acronyms

ACK	Acknowledgment
ACH	Access Feedback Channel
AP	Access Point
ACK	Acknowledgment
ACH	Access Feedback Channel
AP	Access Point
ARQ	Automatic Repeat Request
BCH	Broadcast Channel
CBR	Constant Bit Rate
CC	Central Controller
CDMA	Code Division Multiple Access
C/I	Channel to Interference Ratio
CL	Convergence Layer
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
DCF	Distributed Coordination Function
DIFS	DCF Inter-Frame Spacing
DLC	Data Link Control
EIRP	Equivalent Isotropic Radiated Power
FCH	Frame Control Channel
FDMA	Frequency Division Multiple Access
HCF	Hybrid Coordination Function
HF	High Frequency
HSDPA	High-Speed Downlink Packet Access
HTTP	Hypertext Transfer Protocol
LAN	Local Area Network

LOS	Line of Sight
QoS	Quality of Service
MAC	Medium Access Control
MT	Mobile Terminal
OFDM	Orthogonal Frequency Division Multiplexing
PCF	Point Coordination Function
PER	Packet Error Rate
PIFS	PCF Inter-Frame Spacing
PHY	Physical Layer
RCH	Random Access Channel
RSS	Received Signal Strength
RTS	Request To Send
SIFS	Short Inter-Frame Spacing
SMS	Short Message Service
SNR	Signal to Noise Ratio
SINR	Signal to Interference Noise Ratio
TDMA	Time Division Multiple Access
UMTS	Universal Mobile Telephone System
VBR	Variable Bit Rate
WCDMA	Wideband Code Division Medium Access
WLAN	Wireless Local Area Network

Chapter 1

Introduction

1.1 Motivation

The proliferation of mobile computing and communication devices such as cell phones, laptops, handheld digital devices have given an impetus to bring changes in the todays information technology. The present move is from personal computing to ubiquitous computing in which a user utilizes, at the same time, several electronic platforms through which the user can access all the required information whenever and wherever needed [16]. The nature of ubiquitous devices makes *wireless networks* the easiest solution for their interconnection and, as a consequence the wireless arena has been experiencing enormous growth in the past few years. Untethered from conventional network connections, users can now use their cellular phone to check e-mail, browse the Internet; travelers with laptops can surf from airports, rail stations and other public locations; researchers can exchange files and information by connecting laptops via wireless LANs while attending conferences; at home, users can synchronize data and transfer files between portable devices and desktops, etc.

With this growing demand, the WLANs should satisfy the similar requirements typical of any LAN, including high capacity, full connectivity with the connected terminals and broadcast capability. In mobile telecommunication systems, for e.g., the GSM network offers similar mobility services but it has some capacity and coverage constraints when addressing the residential market. Even with the latest developments and experimentations in the GSM cellular networks, it is shown that there are still places where the present communication platform would fail to provide successful communications [1]. The 3G technologies such as Wideband Code Division Multiple Access (WCDMA), Universal Mobile Telephone System (UMTS), High-Speed Downlink Packet Access (HSDPA),etc. are, however, aimed to achieve better spectral efficiency, higher data rate and a better QoS support. Hence, to meet the demand, WLANs should also be designed to face some issues specific to the wireless environments, such as high data rate, power consumption, security on the air, mobility and bandwidth limitation of the air interface [17]. Evidently the research community faces formidable technical challenges in providing a reliable wireless communication system and network that provide efficient communication platform.

Historically, the infrastructure-based cellular network architecture has been a standard for wireless and mobile communications. In such a cellular network, mobile terminals communicate with the central base station and the base station establishes global connectivity through the infrastructure network. The available bandwidth for the network is shared among users based on the frequency reuse principle. When more terminals are squeezed into the network, there will be terminals who transmit simultaneously in the same frequency band causing interference in the network. Moreover, if there is any obstruction between a sender terminal and the corresponding base station, the transmitted signal will be attenuated to the extent that it cannot be detected by the receiver. There are also situations where users fail to have connectivity in the network. Inevitably, the spectral efficiency and the maximum achievable throughput of the network is limited due to the interference, noise, etc. Therefore, providing higher system capacity, efficient use of spectrum and increased network coverage is a real challenge.

Recently, the ad hoc network architecture has become another wireless communication paradigm for civilian usage. In this network, a collection of wireless terminals form a temporary connection without the need of any fixed infrastructure. A source terminal needs a valid multi-hop relay connection to the destination to successfully send its traffic. The mobile ad hoc network architecture has attractive features that distinguish it from the classical cellular architecture: self-organizing, infrastructure-free terminal to terminal connection within the local ad hoc network, and dynamic adaptation to the fast-changing wireless and mobile environment.

Adding a relaying capability to the cellular network is assumed to potentially increase spectral efficiency, system capacity and service coverage. The system will become an enhanced cellular system with alternative relaying capability to communicate with the base station. Mobile terminals can potentially benefit from the alternative multi-hop path; this could mean greater throughput, larger system coverage or better quality of service (QoS). Similarly, replacing long range high-power transmissions with several short range low-power relay transmissions could reduce energy consumption of the mobile terminals [104, 122, 126]. Other potential benefits from relaying in cellular networks include fast deployment period, fewer infrastructures requirement and peak power consumption reduction.

In short, through augmentation of the relaying capability of ad hoc network with the traditional infrastructure-based cellular network, cost-effective high data rate coverage and throughput can be possible.

1.2 Problem Statement and Thesis Contribution

The demand for wireless networks has increased enormously in the last few years as they offer great flexibility and mobility. One of the most important performance measure of wireless networks is the system capacity. High system capacity is required to satisfy the increase in demand of wireless networks. But since the system capacity cannot be increased arbitrarily, methods of increasing capacity are an essential area of research.

One of the methods of increasing the system capacity is increasing the amount of radio spectrum to be used in the wireless network. However, since radio spectrum is a very scare resource, the solution turns out to be very expensive and changes in spectrum assignment have to be approved worldwide, which often takes years [84].

Use of higher order modulation formats that require less bandwidth and that are more resistant to interference is another possibility. High modulation allows users to use high data rates to send. Equivalently, it allows large number of users to communicate at some constant data rate. The shortcoming of higher order modulation is that it is error prone to noise and interference [61] and it may require a large frequency reuse distance.

Use of sector antennas is another alternative to increase the capacity of a cellular network. The cell is divided into sectors and each sector is served by a sector antenna. Similarly, directional antennas are also used to reduce interference, time dispersion, combat fading and eventually increase the system capacity. Nonetheless, the gain is at the cost of higher infrastructure investment [34, 121, 128].

Recently, the use of relaying attracts many researchers being a potential means to improve the bottle necks of the existing wireless technologies with regard to improving the capacity, coverage and flexibility of wireless networks. As indicated earlier, the multi-hop cellular wireless network, resulting from integration of the relaying feature of mobile ad hoc networks with the traditional infrastructure-based cellular network, has a great promise to improve the capacity of a cellular wireless network. Intuitively, one would expect an increase in capacity by introducing relaying as individual transmissions cover shorter distances, allowing concurrent transmissions which eventually increase the total number of packets delivered per unit time. The focus of this thesis is also to systematically study this problem.

Theoretical studies have shown that using a set of idealized assumptions, the total capacity of the wireless network does indeed increase with the number of nodes, but only very slowly and the throughput per node actually decreases as more terminals are added to the network. And the achievable gain is dependent on the likely locality of communication, i.e., when terminals intend to communicate with their vicinity.

Various attempt are done to apply the theoretical results to a practical multi-hop system. Yet, there exists no overall solution and it is still an open research issue. As indicated in Chapter 4, different power controlling mechanisms are used to reduce interference in multi-hop network, medium access control (MAC) protocols are designed to effectively arbitrate medium access, load aware scheduling and retransmission mechanism are suggested to improve the capacity of the wireless network. Furthermore, the attempts are essentially based on only one or two parameters, i.e., either power control or rate adaptation, routing or scheduling, etc.

The novel approach contributed in this thesis is an integrated approach which targets the transmission rate and power adaptation, the routing, the fair scheduling and multi-channel reuse. First, a routing decision is made if multi-hop communication is worth at all. After the decision, an optimal route is selected to relay traffic in the cellular network and the pertinent transmission power and rate is assigned. Two ways of fairly sharing the system resources are also suggested which reflect both the user and system provider perspective. Based on these fairness schemes, multi-channel scheduling is done which in the end brings a considerable capacity improvement in the network.

There are various issues to be raised when studying the feasibility of multi-hop cellular networks. The following questions give highlights of these issues which are also going to be treated in this thesis:

- How are relaying terminals selected?
- How flexible are terminals to choose the optimal data rate?
- How flexible are terminals to choose the optimal transmission power?
- How is the resource allocation (the channel and communication time) mechanism in the network?
- How are transmissions in the network coordinated for best use of the system resource?
- How much performance gain is obtained by the use of relaying?
- What are the implementation issues and how is the complexity of the system?
- What is the effect of energy efficiency with respect to performance gain in capacity?

1.3 Thesis Structure

The present section briefly summarizes the whole content of the thesis. After this introduction, the remainder of the thesis is organized as follows: Chapter 2 gives background information about the characteristics of wireless environments and resource sharing mechanisms. Among the existing wireless technologies, the IEEE 802.11 and HiperLAN/2 standards are presented as well.

Chapter 3 describes the system model of the multi-hop cellular network to be studied. It includes the different system entities such as the cell, access point and terminal, system parameters. It also states the assumptions to be taken and the metrics to evaluate the system performance.

Chapter 4 presents related work which addresses wireless network capacity issues. In addition to the theoretical study, different approaches such as power control, use of multi-rate and multi-channel, routing and fair scheduling are discussed in relation to improving the network capacity. Finally a summary of the chapter and highlights of the succeeding chapters is given.

Following this, Chapter 5 presents two analytical models: transmission power limited and interference limited systems, to estimate the capacity that can be achieved by introducing relaying in wireless cellular network. An analytical estimation of the probability of finding such relaying terminals in a network is also presented in the end.

Chapter 6 describes the novel approaches used in the thesis to improve the multi-hop cellular network capacity. The approaches are then evaluated by the simulation model described in Chapter 7. The corresponding simulation results and discussions are presented in Chapter 8.

In Chapter 9 the performance of the multi-hop cellular network is evaluated when mobility is introduced in the system. An analytical evaluation is also given to see the extent to which mobility helps to further improve the network capacity. Apart from network capacity, energy efficiency is also a critical issue in wireless networks. Thus, Chapter 10 takes a look at the effects of the proposed mechanism in the total power consumption of the network. Finally, Chapter 11 sums up the whole discussion, points out open issues and future works.

Chapter 2

Background

2.1 Introduction

In the last few years, the use of wireless technologies in the LAN environment has become more and more important, and it is easy to foresee that the wireless LANs (WLANs) will be the solutions for home and office automation as they offer seamless connectivity, great flexibility and mobility compared to the wired LANs [17]. Unlike their wired counterpart, WLANs are typically restricted to a communication range of a single building, a campus, etc. and are usually operated by individuals, not by large-scale network providers [100].

A WLAN should satisfy the same requirements typical of any LAN, including high capacity and full connectivity among terminals. However, to meet these objectives, WLANs should be designed to face issues specific to the wireless environment, such as power consumption, bandwidth limitation of the air interface, mobility and security [17]. Thus, this chapter gives background information about the characteristics of the wireless environment, the existing technologies and the wireless resource sharing mechanisms.

2.2 Channel Model

Mobile wireless systems operate under harsh and challenging channel conditions. Unlike the wired channel, the characteristics of the channel may change rapidly and the behavior becomes unpredictable. There are various factors which affect the propagation channel and cause attenuation in the transmitted signal. These factors are basically due to the propagation behavior of radio signals. The radio signals are propagated according to three mechanisms: reflection, diffraction and scattering. As a result of these mechanism, the propagation channel is characterized by three nearly independent phenomena: pathloss, shadowing and fading [3, 106]. In this section pathloss and fading are briefly discussed.

2.2.1 Pathloss

The phenomena that cause power attenuation in propagated radio signals due to distance and frequency variation is known as pathloss. For two antennas assuming line of sight (LOS) at a distance d, the received power P_{rx} at free space is given by:

$$P_{rx} = P_{tx} G_{tx} G_{rx} \left(\frac{\lambda}{4\pi d} \right)^2 \tag{2.1}$$

where P_{tx} is the transmitted power, G_{tx} and G_{rx} are the transmitter and receiver antenna gains and λ is the wavelength. The factor $\frac{\lambda}{4\pi d}$ is known as free space pathloss factor. The received power in dB is:

$$P_{rx}[dB] = 10 \log P_{tx} + 20 \log \left(\frac{\lambda}{4\pi d} \right) + 10 \log (G_{tx}) + 10 \log (G_{rx}) \tag{2.2}$$

However the signals in land mobile radio do not experience free space propagation. A more appropriate model assumes propagation over a flat reflecting surface (earth) as shown in Figure 2.1. The model is also called two-ray model. The received power then becomes:

$$P_{rx} = 4 P_{tx} \left(\frac{\lambda}{4\pi d} \right)^2 G_{tx} G_{rx} \sin^2 \left(\frac{2\pi h_{tx} h_{rx}}{\lambda d} \right) \tag{2.3}$$

where h_{tx} and h_{rx} are heights of the transmitting and receiving antennas, respectively. Assuming that $d \gg h_{tx} h_{rx}$ and with the approximation $\sin x \approx x$ for small x, the received power reduces to:

$$P_{rx} = P_{tx} G_{tx} G_{rx} \left(\frac{h_{tx} h_{rx}}{d^2} \right)^2 \tag{2.4}$$

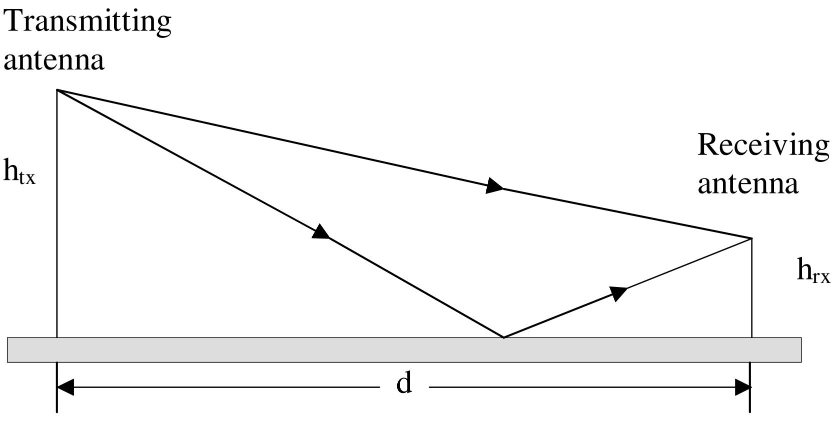


Figure 2.1: A two-ray model for radio propagation over a flat reflecting surface

Thus, the received power in flat reflecting surface decays with fourth power rather than with square and it is not anymore frequency dependent, unlike the free space pathloss. But propagation does not always happen over a flat surface. In realistic channels, the decaying factor depends on the surrounding environment such as urban, sub-urban, and indoor, etc., with values between $1.5 < n < 5.5$. This decaying factor n is also called pathloss coefficient α. For e.g., for a typical urban environment, the pathloss coefficient ranges from 3 to 4 [3, 84, 106].

2.2.2 Fading

Fading is the interference of many scattered signals arriving at an antenna. It is responsible for the most rapid and violent changes of the signal strength and phase. These signal variations are experienced mostly on a small time scale, depending on the velocity of the receiver. The short-term fluctuation in the signal amplitude caused by local multi-path is called small-scale fading, whereas the long-term variation in the mean signal level is called large-scale fading.

Small-scale fading can further be classified as flat or frequency selective, and slow or fast fading [98]. In flat fading the received signal has amplitude fluctuation due to the variation in channel gain over the time caused by the multi-path but the spectral characteristic of the transmitted signal remains unchanged. In frequency selective fading, the received signal is distorted and dispersed due to multiple versions of the transmitted signal, attenuated and delayed in time. This causes time dispersion of the transmitted symbols within the channel which results in intersymbol interference.

When there is relative motion between the transmitter and receiver, Doppler spread is introduced in the received signal spectrum, causing frequency dispersion. If the Doppler spread is significant relative to the bandwidth of the transmitted signal, the received signal is said to undergo fast fading. On the other hand, if the Doppler spread of the channel is much less than the bandwidth of the baseband signal, the signal is said to undergo slow fading.

2.3 WLAN Technologies

The implementation of WLAN follows two different approaches: an *infrastructure-based* approach or an *ad hoc networking* approach as shown in Figure 2.2. An infrastructure-based architecture imposes the existence of a centralized controller for each cell, often referred to as the *access point*. The access point is normally connected to the wired network, thus providing Internet access to mobile devices. In contrast, an ad hoc is a peer-to-peer network formed by a set of terminals within the range of each other that dynamically configure themselves to set up a temporary network. In the ad hoc configuration, no fixed controller is required, but a controller is dynamically elected among all the stations participating in the communication.

The predominantly used transmission technology in WLANs for data communication is the radio wave transmission. Radio waves broadcast on a given frequency can be picked up by any receiver within range tuned to that same frequency. Effective or usable range depends on signal power, distance and interference from intervening objects or other signals [11]. Advantages of radio transmission includes larger coverage area and the capability to penetrate walls, furniture, plants, etc. and

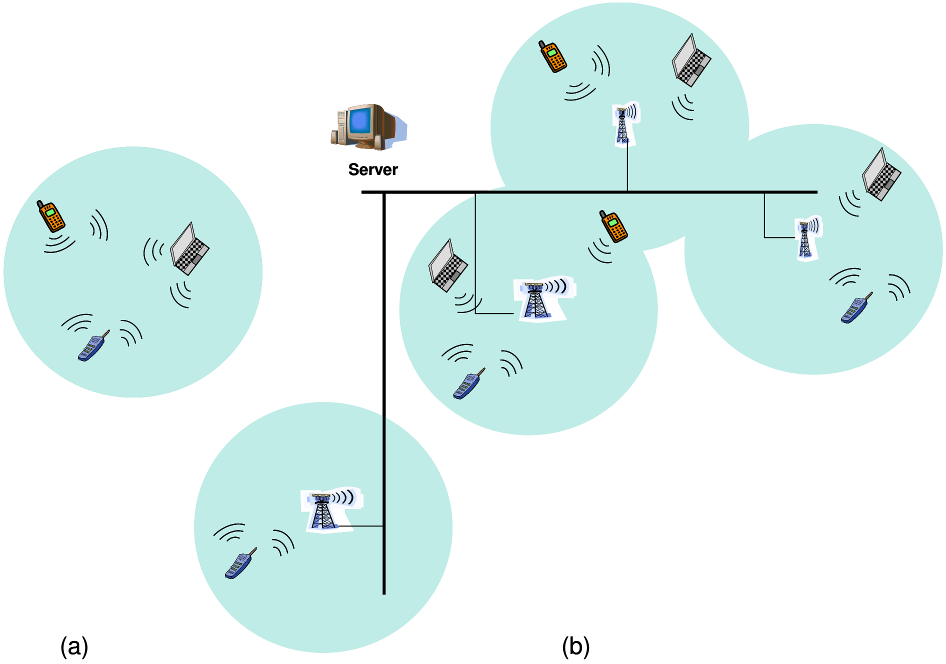


Figure 2.2: WLAN Configuration (a) ad hoc networking; (b) infrastructure-based.

thus requires no LOS if the frequencies are not too high. Furthermore, current radio-based products offer higher transmission rate than infrared. From the most common WLAN technologies, IEEE 802.11 offers both infrared and radio transmission whereas HiperLAN/2 and Bluetooth relay on radio only [100]. In this section, the IEEE 802.11 and HiperLAN/2 technologies are further discussed.

2.3.1 IEEE 802.11

The IEEE standard 802.11 specifies one of the different families of WLANs [16, 84, 100]. The base standard originally allowed data transmission of up to 2 Mbps. Over time, this standard has been enhanced and the extensions are recognized by the addition of a letter to the original 802.11 standard, such as 802.11a, 802.11b, etc. The IEEE 802.11b is a standard for WLAN operation in 2.4 GHz band, with data rates upto 11 Mbps. The IEEE 802.11a is for operation in the 5 GHz band, with data rates upto 54 Mbps. The other families such as 802.11e attempts to enhance the MAC with QoS features to support voice and video over 802.11 network and 802.11g is aimed to have higher speed extensions to the 802.11b.

The IEEE 802.11 standard defines two operational modes for WLANs:*infrastructure-based* and *ad hoc*. Infrastructure mode resembles cellular infrastructure-based network. It is the mode commonly used to provide wireless access to the Internet. In the ad hoc mode, any station that is within the transmission range of any other can start communicating after a synchronization phase. No AP is required, but if one of the terminals operating in the ad hoc mode has a connection to a wired network, terminals forming the ad hoc network gain wireless access to the Internet.

The IEEE 802.11 standard specifies a MAC layer and a Physical Layer for WLANs as shown in Figure 2.3. The PHY layer uses either direct sequence spread spectrum, frequency-hopping spread spectrum, or infrared (IR) pulse position modulation to transmit data between nodes. Infrared is more secure to eavesdropping, because IR transmissions require absolute line-of-sight, contrary to radio frequency transmission, which can penetrate walls and be intercepted by third parities unknowingly.

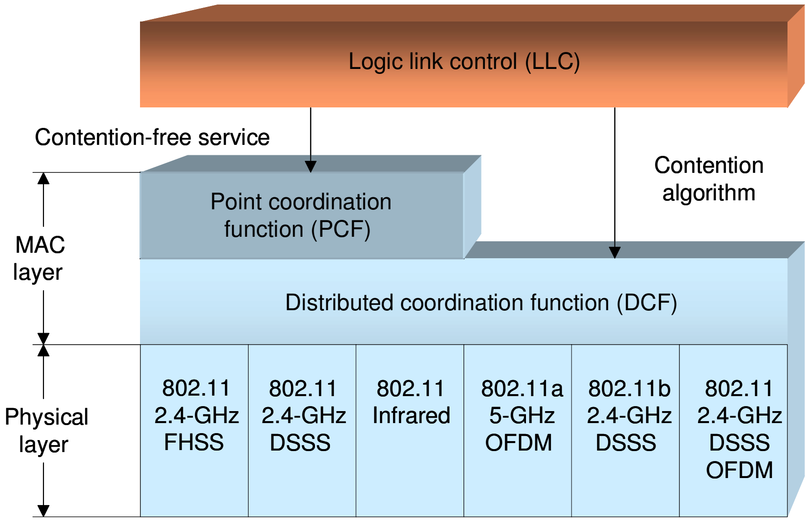


Figure 2.3: IEEE 802.11 protocol architecture

As indicated in Figure 2.3, the PHY layer of IEEE 802.11a is different from the other families as it uses OFDM rather than a spread spectrum scheme. Like in the HiperLAN/2 standard, the IEEE 802.11a operates in 5 GHz band and with variable bit rates 6, 9, 12, 18, 24, 36, 48 and 54 Mbit/s.

The MAC layer offers two different types of services: a contention-based service provided by the *Distributed Coordination Function (DCF)* and a contention-free service provided by the *Point Coordination Function (PCF)*. For both access schemes, the three parameters shown in Figure 2.4 control the waiting time before the medium access. The medium can be busy or idle. The medium is busy either due to data frames or other control frames. During the contention phase several nodes try to access the medium. The DCF inter-frame spacing (DIFS) denotes the longest waiting time and thus the lowest priority for medium access. PCF inter-frame spacing (PIFS) is used for time bounded services and is with medium priority between DIFS and SIFS. The short inter-frame spacing (SIFS) is the shortest waiting time for medium access with highest priority and is defined for short control messages, such as acknowledgments for data packets or polling responses.

The PCF is implemented on top of DCF and is based on a polling scheme. It uses a point coordinator that cyclically polls stations, giving them opportunity to transmit. The DCF provides the basic access method of the 802.11 MAC protocol and is based on carrier sense multiple access with collision avoidance (CSMA/CA) scheme. According to this scheme, when a node receives a packet to be transmitted, it first listens to the channel to ensure no other node is transmitting. If the channel is idle, for at least the duration of DIFS, it transmits the packet. Otherwise, it chooses a random backoff

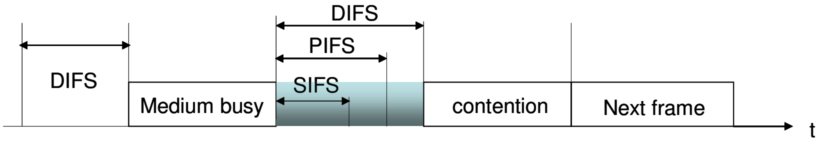


Figure 2.4: IEEE 802.11 medium access and inter-frame space [100]

value which determines the amount of time the node must wait until it is allowed to transmit its packet. During the periods in which the channel is idle, the node decrements its backoff counter. When the backoff counter reaches zero, the node transmits the packet. Figure 2.5 shows this mechanism.

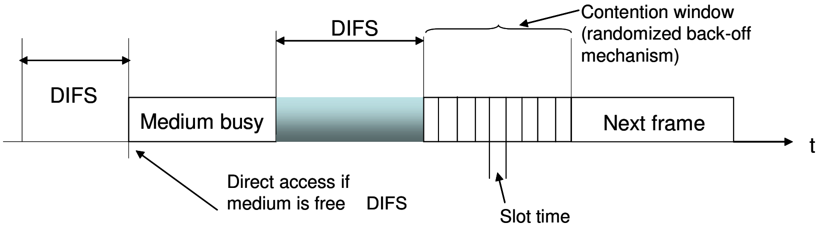


Figure 2.5: IEEE 802.11 contention window and wait time [100]

Usually there is one antenna in a wireless network for both sending and receiving and due to the transceiver hardware, the terminals are not able to listen while sending. For this reason, in the CSMA/CA scheme there is no collision detection capability. Acknowledgment packets (ACK) are sent from the receiver to the sender to confirm that the packets have been correctly received. Colliding station always complete their transmission severely reducing channel utilization as well as throughput, thus presenting challenges to conventional CSMA/CA-based MAC protocols [16].

In wireless ad hoc networks that rely on carrier-sensing random access protocols, such at IEEE 802.11, the wireless medium characteristics face phenomena known as hidden-terminal and exposed terminal problems. The hidden-terminal problem occurs when two or more terminals, say A and C are outside of each other's transmission range and cannot detect each other's transmissions but their transmission ranges are not disjoint. As shown in Figure 2.6, a collision may occur, for e.g., when terminals A and C start transmitting towards the same receiver, terminal B.

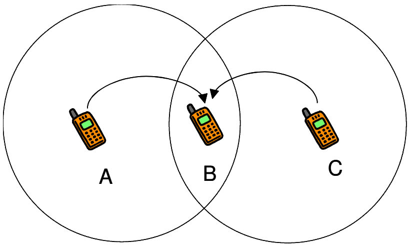


Figure 2.6: Hidden-terminal problem

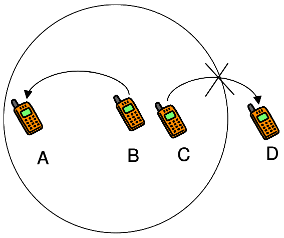


Figure 2.7: Exposed-terminal problem

A virtual carrier-sensing mechanism based on the RTS/CTS mechanism has been included in the IEEE 802.11 standard to alleviate this hidden-terminal problem. The mechanism uses two controls frames, Request To Send (RTS) and Clear To Send (CTS) before the data tranmission. The source terminal sends RTS to the receiving terminal, announcing the upcoming frame transmission. Upon receiving the RTS, the receiver replies by a CTS to indicate that it is ready to receive the data frame as depicted in Figure 2.8. Both RTS and CTS frames contain the total duration of the transmission, i.e., the overall time interval needed to transmit the data frame and the related ACK. This information can be read by any terminal within the tranmission range of either the source or the destination terminal. Hence terminals become aware of transmission from hidden terminals, and the length of the time that the channel is busy for this tranmission.

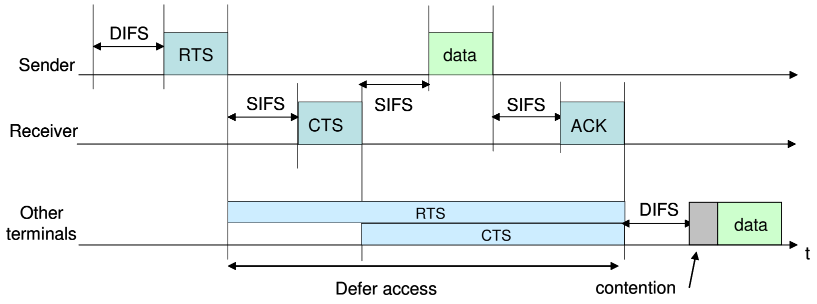


Figure 2.8: IEEE 802.11 hidden-terminal provision for collision free access [100]

The exposed-terminal problem occurs when a permissible transmission from a sender to another station has to be delayed due to another transmission activity between two other mobile terminals within the sender’s transmission range. Figure 2.7 depicts this scenario. In general, such situations result in loss of throughput.

QoS Enhancement for IEEE 802.11 WLANs

QoS is the ability of a network element to provide some levels of assurance for consistent network data delivery [89]. When bandwidth is not scarce, such as in wired LANs, QoS issues are not so important. However, since WLANs have a higher bit error rate, a higher delay and lower bandwidth than a wired LAN, they are originally designed for best effort, low data rate applications. Characteristics of wireless channels make high data rate transmission difficult to achieve: The bit error rate is more than three orders of magnitude higher than in wired LANs. Moreover, high collision rates provoking frequent retransmissions cause unpredictable delays and jitters, which degrade the quality of real-time voice and video transmissions [90].

Different kinds of QoS enhancement schemes for both infrastructure and ad hoc modes have been proposed for 802.11. At the MAC layer IEEE 802.11e is the first supplement to enhance the QoS performance of IEEE 802.11 WLAN. It defines a function called *hybrid coordination function* (HCF) which improves the performance limitations of both DCF and PCF. The 802.11e is assumed to be the future promising framework and is expected to become a new industrial standard [88].

2.3.2 HiperLAN/2

The HiperLAN/2 is a WLAN system standardized by ETSI for 5.2GHz range, giving high data rates to end-users in hot-spot areas. It is mostly aimed at low mobility scenarios such as in offices, homes, exhibition halls, airports, train stations, etc. The HiperLAN/2 standard defines physical (PHY), data-link-control (DLC) and a set of core network specific convergence layer (CL) which can be used with a variety of core networks. Figure 2.9 shows the HiperLAN/2 protocol architecture [27, 29, 30, 61, 93].

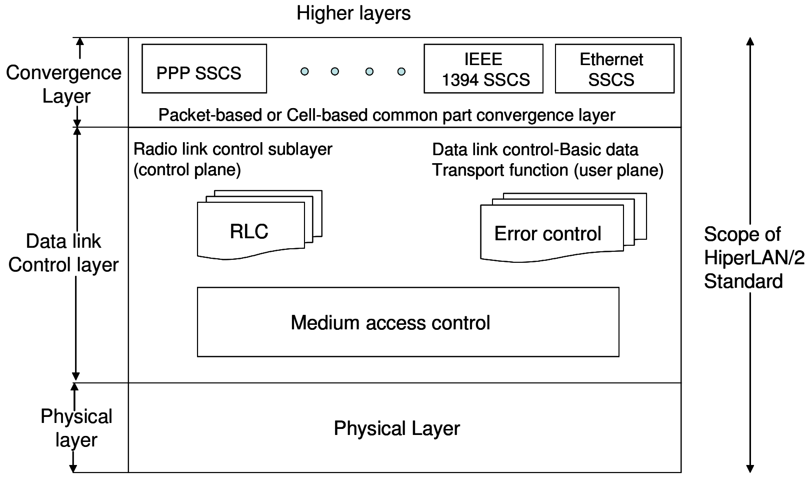


Figure 2.9: HiperLAN/2 protocol architecture

The physical layer (PHY) is responsible for modulation/demodulation of a radio carrier with a bit stream, for synchronization between transmitters and receivers, as well as for forwarding error correction mechanisms, measuring the signal strength and channel sensing. Orthogonal frequency division multiplexing (OFDM) is used as a modulation scheme for HiperLAN/2 due to good performance in highly dispersive channels. This enables HiperLAN/2 to support higher bit rates. Table 2.1 shows PHY layer modes of HiperLAN/2. The PHY modes of IEEE 802.11a are also shown in Table 2.2 for comparison.

PHY Mode	Physical layer bit rate
BPSK 1/2	6 MBit/s
BPSK 3/4	9 MBit/s
QPSK 1/2	12 MBit/s
QPSK 3/4	18 MBit/s
16QAM 9/16	27 MBit/s
16QAM 3/4	36 MBit/s
64QAM 3/4	54 MBit/s

Table 2.1: Physical layer modes of HiperLAN/2 [61]

PHY Mode	Physical layer bit rate
BPSK 1/2	6 MBit/s
BPSK 3/4	9 MBit/s
QPSK 1/2	12 MBit/s
QPSK 3/4	18 MBit/s
16QAM 1/2	24 MBit/s
16QAM 3/4	36 MBit/s
64QAM 2/3	48 MBit/s
64QAM 3/4	54 MBit/s

Table 2.2: Physical layer modes of IEEE 802.11a [84]

One of the design goals of HiperLAN/2 is to maximize the utilization of radio channels. The main mechanism to achieve this goal is by the use of a priori scheduled medium access. This feature makes it different from IEEE 802.11b. Among a set of stations, also called radio cells, one station is declared to be the central controller (CC). Any station that wants to communicate with another station has to announce this to CC, which then allocates time slots in a periodically repeated frame by this station. That means, HiperLAN/2 uses a connection oriented, centrally scheduled time division multiple access (TDMA) to organize the medium access. The main benefits of such organization are to get collision-free data traffic and to straightforwardly support priorities or QoS requirements.

A typical HiperLAN/2 transmission consists of the following steps:

1. A terminal that has data to send announces this to the CC by transmitting a query packet (Resource Request). It also informs the CC the number of packets it wants to send.
2. The CC assigns time slots for this node (Resource Grant) in one of the following MAC frames.
3. Finally the node is allowed to exclusively send in the allocated slots.

The centralized architecture allows the CC to plan all transmissions in its cell. This planning, also called scheduling, is done on the basis of a MAC frame. A MAC frame has a fixed duration of 2 ms and comprises fields of broadcast, frame and access feedback controls, data transmission in the down- and up-link and random access, as shown in Figure 2.10. The duration of broadcast control is fixed, whereas the duration of the other fields is dynamically adapted to the traffic situation.

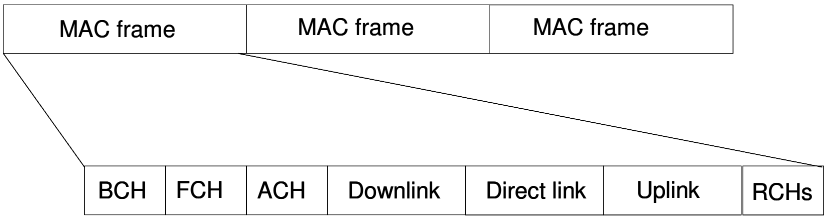


Figure 2.10: HiperLAN/2 MAC frame structure

The broadcast channel (BCH), which contains administrative information the CC broadcasts every MAC frame, mainly enables the control of radio resources. The frame control channel (FCH) contains an exact description of the resource allocation within the current MAC frame. It determines transmission order and duration for all nodes. The FCH conveys the resource grant (RG), each describing one particular transmission for one node while one transmission may contain many packets. The access feedback channel (ACH) conveys information on previous attmepts at random access. Down- or up-link traffic consists of data to or from mobile terminals.

As indicated earlier, the main advantage of a centrally scheduled MAC is the avoidance of collision for any data packets. However, some control packets may collide. If a node has to transmit data, it communicates with the CC using resource request packets. Such packets cannot be scheduled a priori. This problem can be solved in two ways: resource requests can be sent during the Random

Access Channel (RCH) phase of HiperLAN/2 MAC frame, which may collide with other similar requests. Such mechanism is defined in the standard to notify stations about the success of their RCH access [28]. Alternatively, the CC of a cell can periodically assign uplink transmission slots for resource request packets to each node, giving them the possibility to request new resources. In such a case, there are no collisions in the MAC frame, but at the expense that this permanent polling may produce a large number of control packets depending on the total number of nodes in the cell [24].

The FCH-based MAC frame organization ensures that all devices within the cell know about all MAC frame data transfers as they all receive the FCH. The FCH structure opens up possibilities to realize various MAC-based improvements of wireless communication, like the Direct Link Mode as defined in the Home Extension of the HiperLAN/2 standard [31]. The extension allows any two nodes being within each other's radio range to directly exchange data, under the control of CC. With this possibility the HiperLAN/2 protocol is also extended to realize multi-hop relay communication as in [50]. In this thesis, this extended relaying feature is exploited to optimize the network capacity.

2.4 Multi-hop Systems

The provision of high capacity, low cost and reliable wireless communication continues to be a challenging aspect of future wireless communication system. In this sense, the infrastructured and ad hoc networks play a significant role in wireless communication systems. These systems have the potential to improve the capacity and/or extend the coverage cost-effectively [76, 119, 120].

The limited communication range, however, introduces some difficulties in wireless communication. For example, if terminals spread over a wider region, the distance between terminals will be too large to communicate with each other using single wireless transmission. One possible recourse is to use other terminal devices to help in communication between distant terminals: These terminals would volunteer to relay data from one terminal towards its destination; data would travel over several wireless communication hops, making such a network a *wireless multi-hop* [59]. Multi-Hop networks are thus mobile communication structures dynamic in nature and consist of mobile terminals with retransmission capability on behalf of other mobile terminals. The main reasons for introducing multi-hop networks in 3G systems are to overcome the capacity and coverage problems of CDMA/TDMA systems, and thus increase and improve the services for users, as well as to reduce the cell planning difficulties for operators [119].

The following are assumed to be operational multi-hop architectures within 3G systems [119]:

1. Relay multi-hop network with infrastructure support, being the relays fixed or mobile. Three main benefits could be provided within this architecture:
    (a) The coverage extension of High Bit Rate services
    (b) Service coverage for those users which are beyond the coverage area of the AP

   (c) The potential interference reduction when any packet connection evaluates the possibility of a multi-hop versus a single hop link.

2. System internetworking for vertical handover, allowing high bit rate coverage of picocells (as the most promising possibility for 4G systems).

3. Ad hoc multi-hop networks (with no infrastructure support), providing connection of close users within picocells.

The foremost benefits of the deployment of a multi-hop network are the reduction in system interference and hence increased network capacity, and extended coverage for high bit rate users and for those users beyond the coverage area of the AP. Moreover, breaking a larger path down into a number of smaller hops reduces the aggregate radio path loss, total transmission power and smoothes the distribution of interference. As a direct consequence, the subscriber's battery life can potentially be increased and the fear from radiation hazards be erased.

2.5 Medium Access

In wireless networks, a certain frequency spectrum is assigned for a specific service or system. Since this spectrum is a scarce resource that cannot easily be extended, the system should make a provision to allow as many communications as possible within the allocated spectrum. The way in which multiple users access a communication medium is typically by one of the methods based on either frequency, time or code allocation. In this section, the most commonly used ones: The time division multiple access (TDMA) and the Code Division Multiple Access (CDMA) are discussed. The orthogonal frequency division multiplexing (OFDM) is also discussed alone with the access schemes.

2.5.1 TDMA

The TDMA scheme is well established and a flexible access scheme in shared radio networks. It allows several users to share the same frequency by dividing it into different time slots. In TDMA, the wireless channel is generally organized in frames, where each frame contains a fixed number of time slots. The mobile nodes negotiate a set of TDMA slots in which they can transmit their traffic. If a centralized controller exists, it is responsible to assign the slots to the nodes. The nodes are active for short periods of time on the same frequency of operation. Besides, temporal guard intervals are used not to overlap the signals from adjacent users [84]. In this way, transmissions are collision free and it is possible to schedule node transmissions according to fairness and QoS criteria. TDMA requires less stringent power control due to reduced inter-user interference. An appropriate scheduling for slot assignment and the absence of collision guarantee bounded delays [16].

However, in a mobile network environment the re-assignment of slots after topology change makes the legacy TDMA scheme very inefficient [16]. Also, guard spaces are needed in between the slot, which reduces the resource utilization efficiency.

2.5.2 CDMA

The CDMA is a method of multiple access that does not divide up the channel by time (as in TDMA), or frequency (as in FDMA), but instead individual transmissions are encoded with pseudo-random digital sequences. This allows terminals to use the full available spectrum. Thus, terminals are active at the same place at the same time without being interrupted. CDMA consistently provides better capacity for voice and data communications than other commercial mobile technologies, allowing more subscribers to connect at any given time.

CDMA is also flexible and requires less planning as all terminals can transmit at any time. However, it needs complex receivers and complicated power control for senders.

2.5.3 OFDM

OFDM is a modulation scheme that is especially used for high data rate transmission in delay dispersive environments. It converts a high rate data stream into a number of low rate streams that are transmitted over parallel, narrowband channels [84]. The fundamental principle of OFDM is the use of overlapping subcarriers to modulate parallel data streams. This makes OFDM much more bandwidth efficient than the conventional non-overlapping multi-carrier techniques. The subcarriers in OFDM system need to be orthogonal, so that they do not interfere each other. OFDM communication systems naturally alleviate the problems of multi-path propagation with its low data rate per sub-carrier as it is only a fraction of conventional single carrier system having the same throughput. It also ensures flat fading condition on each subcarrier [127].

OFDM has received enormous attention mainly due to its two properties: First, OFDM has a higher spectral efficiency compared to traditional Frequency Division Multiple Access (FDMA) systems. In addition to this, OFDM systems are not effected by frequency selective fading channels. In single carrier systems, a single fade or interferer may cause the entire link to fail. However, in a multi-carrier system, such a fade would affect only a few of the subcarriers. As a consequence, there is no need for equalization at the receiver, a costly task in single carrier modulation systems. This advantage is especially significant in the context of wireless communications. To make use of this advantage, the physical layers of the HiperLAN/2 and IEEE 802.11a use OFDM as their underlining radio technology [21, 39].

2.6 Frequency Reuse

The tremendous growth of the wireless users coupled with the bandwidth requirement of multimedia applications require efficient reuse of the scarce radio spectrum. Efficient use of radio spectrum is also important from a cost-of-service point of view, where the number of base stations required to service a given geographical area is an important factor.

A cellular mobile radio system relies upon *frequency reuse*, where users in geographically separated cells simultaneously use the same carrier frequency or channel [106]. To maintain sufficiently high signal-to-interference ratio, the same chanel should not be reused in cells which are too close to each other. Obviously, for large reuse distance, large numbers of channels are required. Consider the hexagonal cells shown in Figure 2.11 to determine the minimum reuse distance. To do that, first proceed i hexagons in the y direction from the origin, then turn $60°$ and proceed k hexagons in the new direction as in the figure. The same channel is to be used in the first and the last cells. The reuse distance D between the two cells is given by [84]:

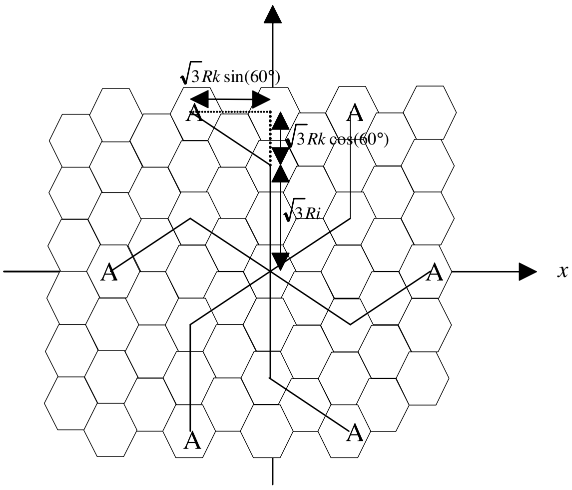


Figure 2.11: Minimum reuse distance [84]

$$\begin{aligned} D &= \sqrt{(\sqrt{3}Ri + \sqrt{3}Rk\cos 60)^2 + (\sqrt{3}Rk\sin 60)^2} \\ &= \sqrt{3}R \cdot \sqrt{i^2 + ik + k^2} \end{aligned} \tag{2.5}$$

where R is the radius of the cell. The task of frequency planning is, thus, to find the values of i and k that make the distance in Equation (2.5) larger than the required reuse distance. In other words, the values minimize the cell cluster size and thus maximize spectral efficiency, while at the same time satisfying the minimum reuse distance. Following Equation (2.5), the relationship between cluster size N and the parameters i and k is [60, 84, 106]:

$$N = i^2 + ik + k^2 \qquad \text{i, k = 0,1,2,...} \tag{2.6}$$

The allowable cluster sizes are $N = 1, 3, 7, 9, 12, 13, 16, 19, 21, ...$ $N = 1$ corresponds to the case where all channels are used in all cells as in CDMA systems and $N = 7$ is a typical value for TDMA system as in GSM. And the ratio D/R is called the *co-channel reuse factor* given by:

$$D/R = \sqrt{3N} \tag{2.7}$$

All cells which use the same channel are referred to as *co-channels*. The basic impediment factor in radio spectrum reuse and that limits capacity of cellular systems is the co-channel interference. Frequency reuse also introduces *adjacent channel interference* which arises when neighboring cells use carrier frequencies that are spectrally adjacent to each other. In this case the power density spectrum of the desired and interfering signals partially overlap. Channel impairments such as cross-talk, premature hand-offs and dropped calls may result from adjacent channel interference.

2.7 Transmission Power and Interference Limited Systems

To get a good insight to the basic capabilities of wireless systems and to properly budget the wireless link, it is essential to study the different system parameters separately [84]. In this section transmission power and interference limited systems are explained.

2.7.1 Transmission Power Limited

In a situation where there is only one access point communicating with a mobile terminal, the received power decreases as the terminal moves further away from the AP. To maintain the signal-to-noise ratio for reliable communication, the terminal is allowed to increase the transmission power. However, the terminal cannot arbitrarily increase the power due to isotropic radiated power limitation and the limits of HF (high frequency) power amplifier . The performance of the system is thus determined by the strength of the useful signal and the noise. Such a system is called *transmission power limited* or *noise limited*. Typical examples of such a system are WLANs with one AP and a cordless phone [84].

2.7.2 Interference Limited

There are situations where interference is very strong that it completely dominates the performance of the wireless system and noise can be neglected. In such a case, the system is called *interference limited*. An example of such a situation is when there are, for e.g., two access points separated by a distance D that operate at the same frequency. The signal-to-interference ratio becomes a useful parameter to determine the performance of the system. Studying such a system is also useful to determine the minimum distance D to guarantee a satisfactory performance for 90% of the time when mobile terminals are at the cell boundaries [84].

Chapter 3

System Model

3.1 Introduction

It has been pointed out earlier that the infrastructure-based cellular network architecture has been the standard for wireless and mobile communications. In such a network the mobile terminals should always communicate with the central base station or the AP either to send/receive data or to have global connectivity.

There are various advantages in such infrastructure-based networks with central controllers. There is a better control of resources such as spectrum and time. The central organization makes it is easier to allocate and manage these resources according to the user need or fairness rules. Necessary network information can be stored in the central base station which can later be used for resource management or other uses. This reduces the complexity and the processing power of terminals available in the network as most of the network functions are concentrated within the central controller.

On the one hand, cellular infrastructure-based networks are less flexible as all terminals within a cell should always communicate with their central base station. Specially far terminals may need to communicate with their base station using highest transmission power. Thus, it makes the communication in cellular networks interference prone which eventually reduces the cellular network capacity and coverage.

On the other hand, due to the absence of infrastructure and base station, resource management in ad hoc networks is difficult. For example, bandwidth allocation to each node changes dynamically according to the mobility scenario. But since the role of a node as host or relay also change together with the node mobility and the dynamic topology, bandwidth control is difficult [43]. Consequently the aggregate network capacity is also time varying. The dynamic topology change also requires large and frequent exchanges of control information among the network nodes. As all updates in the wireless communication environment travel over the air, ad hoc networks are, thus, costly in resource [99].

When the features of mobile ad hoc networks are combined with that of infrastructure-based cellular networks, the multi-hop cellular wireless network is formed. In other words, the system becomes an enhanced cellular system with alternative relaying capability to communicate with the base station. Mobile terminals can benefit from the alternative multi-hop path; this could provide greater throughput or better quality of service (QoS). Similarly, replacing long range high-power transmissions with several short range low-power relay transmissions could reduce energy consumption to the mobile terminals [104, 122, 126]. Other potential benefits from relaying in cellular networks include fast deployment period, fewer infrastructures requirement, peak power consumption reduction. Even from ad hoc mobile networks point of view, the future ad hoc network is towards seamless integration with other wireless networks and the fixed Internet infrastructure appears to be inevitable[16].

It is envisioned that relaying would be practical in urban environment where the user population is dense. Moreover, in realistic scenarios such as in exhibition halls, airport and supermarkets where many users communicate with each other using wireless networks, network capacity is a big concern. In such cases, relaying is an attractive choice. Therefore, for all the reasons discussed, the multi-hop cellular wireless network is chosen as an interesting wireless network and a particular emphasis is given, in this thesis, on how to improve its capacity by virtue of relaying.

The remainder of this chapter is organized as follows. The channel model of the system is explained in Section 3.2. Section 3.3 describes the system entities: the cell, the AP and the mobile terminals. The different system parameters are then discussed in Section 3.4. In Section 3.5 the communication protocol structure is explained. Finally in Sections 3.6 and 3.7 the load model and the system metric are discussed .

3.2 Channel Model

The propagation model considered here is derived due to the micro-cellular characteristic of HiperLAN/2 and it is based on measurements (empirical model) in [61, 91].

Consider a mobile terminal sending its traffic to another mobile terminal or AP at a transmitting power of P_{tx}. The received power P_{rx} is calculated as:

$$P_{receive} = \frac{P_{transmit}}{10^{\frac{1}{10}L}} \tag{3.1}$$

where L is the pathloss. As discussed in Section 3.2, this pathloss is characterized by a two-ray model. For a direct and ground reflected rays, the received power shows inverse square and inverse 4^{th} power decay. The transition distance between the two decay factors is the critical distance as shown in Figure 3.1. The corresponding pathloss is given by:

$$L_1(d) = L_{01}(d_o) + 10 \cdot n_1 \cdot \log_{10}(\frac{d}{d_o}), \quad \text{for } d < d_{crit} \tag{3.2}$$

$$L_2(d) = L_{02}(d_{crit}) + 10 \cdot n_2 \cdot \log_{10}(\frac{d}{d_o}), \quad \text{for } d \geq d_{crit} \tag{3.3}$$

where d denotes the receiver-transmitter separation distance, $L_{01}(d_o)$ is the reference pathloss at the reference distance d_o and $L_{02}(d_{crit})$ denotes the required pathloss for a continuous pathloss curve at d_{crit}.

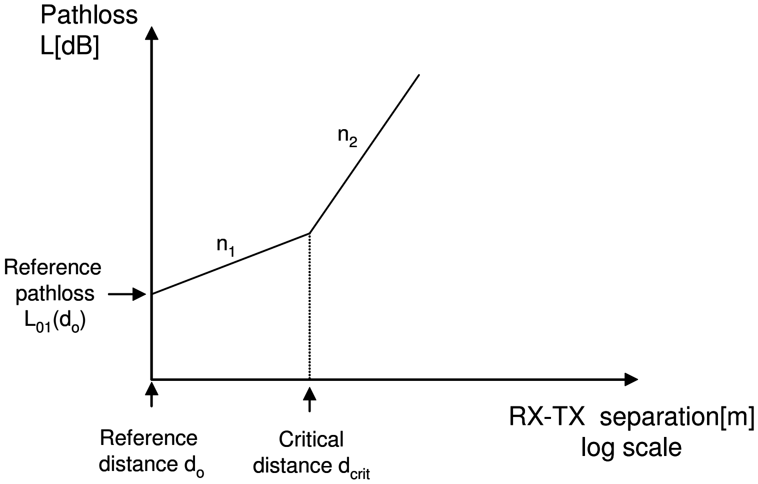


Figure 3.1: Pathloss with two-ray model

Parameter	Default value
$L(d_o)$	46
d_o	1m
n_1	2
n_2	4
d_{crit}	70m

Table 3.1: Initial model parameters [91]

Assuming free space pathloss, inverse square law for $d < d_{crit}$, inverse 4^{th} power law for $d \geq d_{crit}$ and from the initial model parameters shown in Table 3.1, the pathloss becomes

$$L_1(d) = 46 + 10 \cdot 2 \cdot \log_{10}(\frac{d}{m}), \quad \text{for } d < d_{crit} \tag{3.4}$$

$$L_2(d) = (\approx 10) + 10 \cdot 4 \cdot \log_{10}(\frac{d}{m}), \quad \text{for } d \geq d_{crit} \tag{3.5}$$

In realistic scenarios, there is inverse square law decay when the receiver is close to the transmit antenna. But at larger distances, the pathloss coefficient varies depending on the propagation environment [84]. The cellular network that is going to be pondered in this thesis has a cell radius not more than the critical distance, i.e., 70 m. Therefore, it suffice to use $L_1(d)$. With a reference pathloss at 1 m and a varying pathloss coefficient α, the pathloss then becomes:

$$L = 46 + 10 \cdot \alpha \cdot \log_{10} d \tag{3.6}$$

3.3 System Entities

The infrastructure-based multi-hop cellular wireless system to be considered comprises of a number of entities: a number of cells, access points and terminals. The properties of these entities are discussed in this section.

3.3.1 Cell

In an infrastructure-based cellular network, the area served by the network is divided into subareas called cells. Each cell contains an access point and a number of terminals. The access point is placed in the center and serves as the central controller of every communication within the cell. Assuming that the terminals are located on a planar surface having a uniform circular pathloss, the most natural way of representing the cells is with circular shape. However, since circles cannot fill the planar surface without a gap or overlap, tessellated regular hexagons are used to represent the shape of the cells. In a real world, the shape of the cells is actually highly irregular and there may be overlaps or dead spots (regions with no communication) [84, 129].

Pertinent to the resource allocation scheme, the available spectrum is then shared by the users within each cell. Following the discussion in Section 2.6, cluster sizes of $3, 7$ and 19 are typically studied here. The choice of these sets is logical since HiperLAN/2 system, for e.g., provides upto 19 channels for communication, and hence we can be flexible in choosing the number of frequencies to be reused.

Another important issue when studying cellular networks is defining the cell size. The size depends on factors like pathloss coefficient, tranmission rate, traffic load and maximum allowable packet error rate. In practice, defining the cluster size and the number of cells to be used in the communication service area is not an easy task. Also, cell sizes are variable due to cell-breathing – the effect of shrinking or growing of cell size over time depending on the traffic load. But in general, the available frequency spectrum and cost are constraining factors in cell planning [129].

Definition 3.1 *Given that a mobile terminal in a cell is capable of sending its data with a set of* k *discrete transmission rates* $r_1, ..., r_k$*. For a certain maximum target packet error rate PER and pathloss coefficient, each of these discrete transmission rates are associated with a set of maximum transmission radii* $d_1, ..., d_k$*, respectively, where* $d_i < d_j$ *for* $i < j$*. The radius of a cell* R_c *is then:*

$$R_c = max(d_1, ..., d_k) \tag{3.7}$$

For a given PER, terminals with lower transmission rates traverse larger transmission radii than with higher rates [61]. Thus, the size of a cell is defined, in other terms, as the maximum transmission radius that a mobile terminal can traverse when directly communicating with the access point using the lowest transmission rate and the maximum possible transmission power for a certain pathloss coefficient α.

Figure 3.2 shows the relationship between cell radius, maximum target PER and pathloss coefficient for the lowest transmission rate. For typical pathloss coefficients in urban environment [3] which are between 3 and 4, the radius of the cells is in the ranges of about 50m. Thus, the cells considered here are also within these ranges.

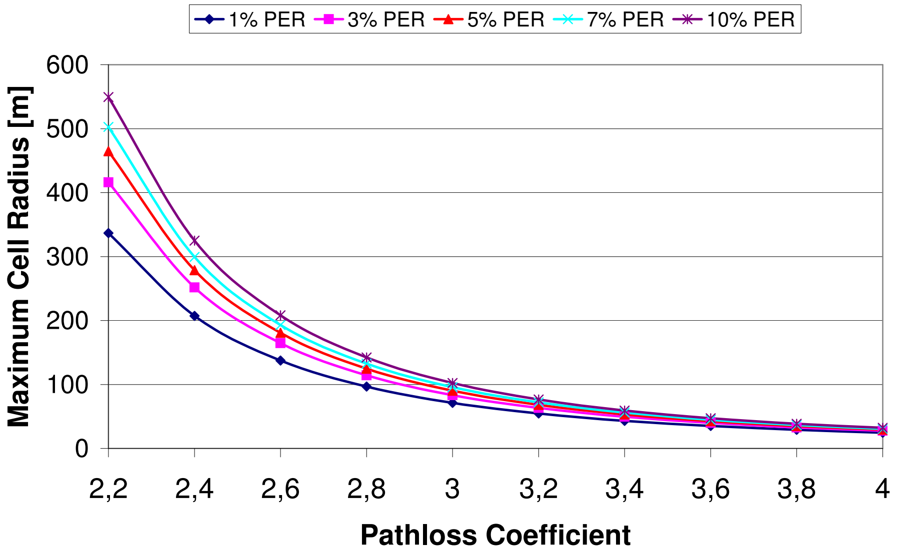


Figure 3.2: Relationship between maximum cell radius, PER and pathloss coefficient with the lowest transmission rate

Corollary 3.1 *Given a set of cells $C_1, ..., C_n$ in a certain service area. If the radius of cell i is r_c^i with $r_c^i \leq R_c$, then all terminals in cell i are reachable by the AP. Conversely, if $r_c^i > r_c$, not all the terminals are directly reachable by the AP and hence there is a need for an intermediate terminal between the AP and the terminal to establish communication.*

Following these, terminals at the periphery of a cell with radius larger than R_c may not always have direct connectivity with the AP even if they use the slowest transmission rate. Multi-hoping is one way to provide connectivity to such terminals and extend the cell coverage. However, coverage extension of cellular networks through multi-hopping is not the focus of this thesis. Instead, multi-hop communication is applied to improve the system capacity of a fixed cell radius, with all its terminals being directly reachable by the AP.

3.3.2 Access Points

The access points in an infrastructure-based cellular network are the central controllers of each cell having various system functionalities. They are located at the center of each cell, connected to the infrastructure network. Thus, they provide global connectivity to the mobile terminals in each cell.

Mobile terminals which are within the range of their AP should always establish a link to their AP in order to communicate either with the AP itself or with other mobile terminals in the cell. The AP is assumed to be the central controller and decision maker in the system, akin to the CC of HiperLAN/2 discussed in Section 2.3.2. It has every information about the terminals in its own cell, especially the channel gain estimates between the terminals. Based on the information, the AP first decides if a mobile terminal needs relaying to send its traffic. Then it assigns the terminal to a particular time slot, transmission power and data rate. In other words, the AP is the *scheduler* of the entire communication within its own cell. It also provides synchronization in the cell.

Furthermore, the AP is assumed to have an omni-directional antenna which does not contribute any loss or gain when communicating with every mobile terminal in its cell. Also it does not have any battery constraint as its is already connected to the infrastructure network.

3.3.3 Terminals

In the traditional infrastructure-based cellular network, terminals are limited either to sending/receiving traffic by directly connecting to the access point only. With the integration of ad hoc capability to the traditional cellular network, mobile terminals not only send/receive traffic but they can assist in relaying as well.

At some instant of time mobile terminals in a cell exist in different modes. Mobile terminals are said to be *active* if they have some traffic to send and require connection. There are also terminals which are *idle*. Idle here refers to a terminal with no traffic of its own but can forward traffic. For e.g., mobile terminals such as laptops with wireless communication capability and plugged to a constant power source may not have traffic to send all the time and may remain idle. But such terminals can assist in multi-hop communication. Hence, the type of terminals considered here are those with their own traffic and terminals with no traffic of their own but are ready to assist in forwarding.

In a real system, there is a cost of resource consumption when terminals relay traffic from other terminals. Eventually, the terminals do not want to incur this cost and may not always be willing to assist multi-hoping. To handle such cases, incentive mechanism for collaboration have been proposed in [18, 23]. The incentive is provided through the concept of a user having a credit balance, which receives an initial endowment when the user joins the network. The credit balance accumulates notional credit accrued by forwarding traffic for other users, while any traffic from a particular user decreases the credit balance based on the cost of forwarding its own traffic to its destination. The amount of traffic that a user can generate is directly related to its current credit balance, which is the user’s incentive to both act as a relay node for other users and move to locations within the network where it can forward more traffic. However, identifying collaborative terminals for multi-hoping by such incentive mechanisms in the cellular network under study is beyond the scope of this thesis. The willingness/readiness of terminals in multi-hoping is the underlining assumption throughout the thesis. But to avoid unfair/unnecessary overloading of relaying terminals, especially for those which are located close to the AP and are highly likely to be selected, the number of supported relay traffic is limited.

Let M be a set of active mobile terminals in a cell. The size of this set varies depending on factors such as the traffic load intensity and the instant of time the observation is made. These factors also affect the distribution of the set. Two types of terminal distribution are considered in the system: uniform distribution and terminals in clustered form. In the former case, terminals are uniformly distributed in the cell, with a constant probability density of active terminals, i.e., for an active terminal located at (X, Y) in a given cell with a total coverage area of A, the probability density is given by:

$$p_{XY}(x,y)dxdy = \Pr\{X \in (x, x+dx), Y \in (y, y+dy)\} = \frac{1}{A}dxdy \tag{3.8}$$

In the latter case, a group of terminals may be either at the vicinity of the AP or at the periphery of the cell forming a clustered distribution. Details on topology generation is discussed in Chapter 7 .

For a given distribution, terminals in a cell are identified by their location vector X and Y, and by their unique MAC-ID. The AP has full knowledge of the terminals within its coverage range. These mobile terminals could be stationary or have a random walk mobility characteristic. In the mobility case, terminals move from their current location to a new location by randomly choosing a direction and speed to travel. The new speed and direction are both chosen from predefined ranges $[v_{min}, v_{max}]$ and $[0, 2\pi]$, respectively [15]. Issues related to mobility are addressed later in Chapter 9.

Unlike the APs, the terminals have constraint in a battery source. The constraint may not necessarily be severe for cases like laptops with wireless capability or mobile terminals in vehiculars with relatively large battery capacity. The terminals are assumed to have an omni-directional antenna with uniform circular radiation loss. Furthermore, they are capable of sending traffic with a set of k discrete transmissions rates. These discrete rates are the different modulation levels offered by the physical layers of HiperLAN/2 and IEEE 802.11a wireless systems discussed earlier in Section 2.3.

3.4 System Parameters

3.4.1 Transmission Power

Transmission power is an important parameter in wireless systems which influences the transmission mode and interference situation. Transmission power cannot simply be increased to achieve a high signal-to-noise ratio or better link quality. Increasing the tranmission power of a terminal will increase the interference on other terminals which transmit in a similar channel. These terminals further increase their tranmission power to combat the increased interference, which further aggravates the interference. Moreover, tranmission power cannot exceed a certain limit due to regulatory limitation on equivalent isotropic radiated power (EIRP). Thus, selecting a transmission power is not a trivial problem in a wireless system.

Suppose that a mobile terminal i is transmitting to mobile terminal j over channel k. The received power at j, $P_{rx,j}$ is given by:

$$P^k_{rx,j} = P^k_{tx,i} G^k_{ij} \tag{3.9}$$

where $P^k_{tx,i}$ is the transmission power selected by terminal i. G_{ij} is the link gain between the transmitter i and receiver j and it is obtained according to the channel model discussed in Section 3.2. For successful transmission $P^k_{tx,i}$ should be a non-negative power that satisfies the minimum target PER and SINR.

3.4.2 Interference

The sole source of interference in the system under study is *co-channel* interference, due to the simultaneous transmissions of terminals that simultaneously use the same channel. It is assumed adjacent channel is fully suppressed and hence there is not adjacent channel interference.

Suppose that there are simultaneous transmissions in co-channel cells in a wireless network. Let T be the set of terminals sending data over channel k, and let $i \in T$ and sends traffic to j. The interference power received at j, $P_{I,j}$, is given by:

$$P^k_{I,j} = \sum_{l \in T, l \neq i} P^k_{tx,l} G^k_{jl} \tag{3.10}$$

where G^k_{jl} is the link gain between the interfering terminal l and the receiving terminal j. Thus, total interference power is the sum of effective power of the individual interference components in the wireless channel.

3.4.3 Signal to Interference and Noise Ratio

A transmission is said to be successful if it satisfies the following condition:

$$\Gamma^k_{ij} = \frac{P^k_{tx,i} G^k_{ij}}{\sum_{l \in T, l \neq i} P^k_{tx,l} G^k_{jl} + N_{jo}} \geq \gamma^k_j \tag{3.11}$$

where Γ^k_{ij} the actual signal to interference and noise ratio when terminal i is sending to j and γ^k_j is the minimum signal-to-noise ratio needed to achieve the target PER for a successful transmission. N_{jo} is the noise power at the receiver terminal j. The ratio of energy per physical layer input bit to noise power density N_o is a good measure for the link quality in the presence of white noise. In this thesis, the γ^k_j values are obtained from the PER to SINR and modulation mapping of the HiperLAN/2 physical layer [61] as shown in Figure 3.3. But since the physical layer of HiperLAN/2 and IEEE 802.11a are similar, it is expected that the results be directly generalizable.

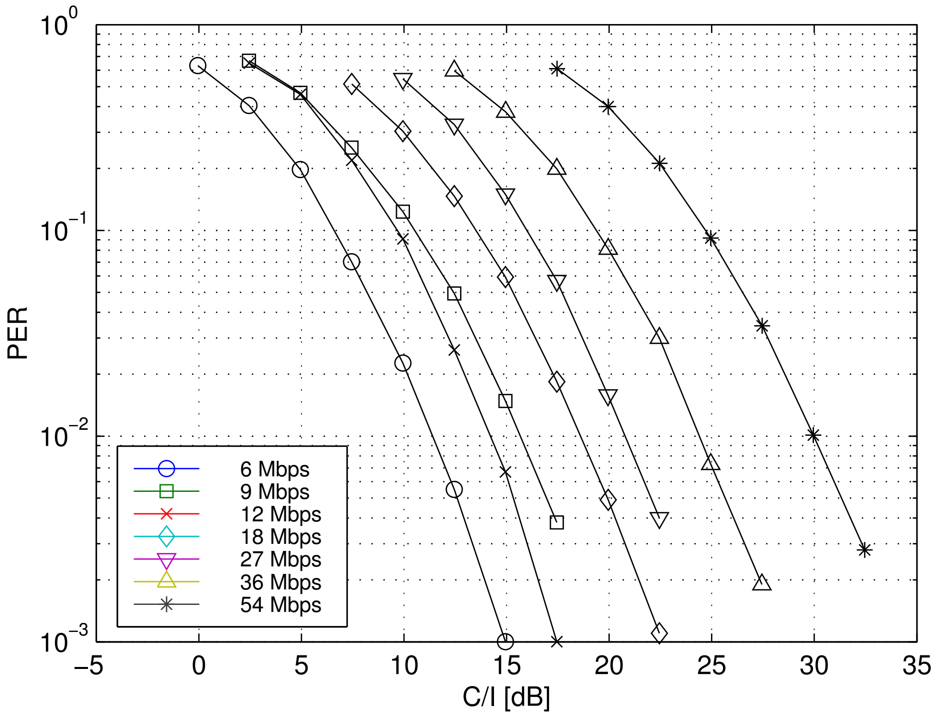


Figure 3.3: PER to modulation and SINR mapping [61]

3.4.4 Receiver Sensitivity

The radio receiver sensitivity is the received signal power level at the antenna reference point of the receiver at which the packet error rate shall be less than or equal to 10% [27]. The sensitivity levels for the seven PHY modes of HiperLAN/2 are listed in Table 3.2. The table also shows the SNR value by setting the received power P_{rx} to the receiver sensitivity, which is approximately the same as the SNR value to a PER of 10%.

PHY mode	P_{rx} = receiver sensitivity(dBm)	white Noise power	SNR	SNR (PER = 10%)
1	-85	-91.87	7.87	≈ 6
2	-83	-91.87	9.87	≈ 10
3	-81	-91.87	11.87	≈ 9
4	-79	-91.87	13.87	≈ 13
5	-75	-91.87	17.87	≈ 15
6	-73	-91.87	19.87	≈ 19
7	-69	-91.87	24.87	≈ 24

Table 3.2: SNR and receiver sensitivity [29]

3.5 Communication Protocol

3.5.1 Physical Layer

To enable rate-adaptive relaying, a physical layer is required that provides several data rates to choose from, trading off speed against communication distance. As a case study, the physical layer of HiperLAN/2 [61] is used here, which provides seven different data rates (modulation plus coding rate). Since this physical layer is very similar to the one employed by IEEE 802.11a, it is expected that these results to be directly generalizable.

3.5.2 MAC Structure

The essential choice for a MAC structure is between a contention-based and a scheduled medium access approach. Contention-based systems are currently more popular, but for the question at hand, the particular medium access control *within* a cell should not play the dominating role. A scheduled access control allows as a finer-grained control over how resources are used and assigned within a cell, admitting deeper insights into the structural properties of our presented concept. TDMA based scheduling, in particular, allows explicit control over how resources are split among and assigned to terminals within a cell, something that is not immediately possible with a CSMA structure.

Therefore, we decided to use a TDMA-based MAC structure to be able to control resource assignments in detail. Like the physical layer, the details of the MAC correspond to HiperLAN/2—a frame of 2 ms, which is split into slots, assigned to a specific sender–receiver combination (which does not have to include the AP).

3.5.3 Central Organization

Commensurate with the choice of scheduled medium access, the per-cell centralized decision making process is opted. The access point only has information about terminals in its own cell—especially, channel gain estimates between terminals—and bases all its decision on this information. In particular, the AP decides in which time slot a given sender and receiver (both belonging to its own cell, including the AP itself) should communicate, which transmission power and which data rate to use. The length of a slot can vary, depending on the chosen data rate. A typical fixed TDMA frame comprising of both up- and down-link communication is shown in Figure 3.4.

Figure 3.4: typical fixed TDMA frame

Channel gain estimates can be obtained by the AP in various ways, e.g., using a concept similar to HiperLAN/2's "radio map" [30] or using approaches like optimistic rate adaptation, e.g. similar to [49] (occasionally switching to a faster rate to check whether communication is still possible and falling back to a slower rate if such an attempt fails).

The advantage of such a centralized organization is a much better control over the system's behavior, giving better insights into the potential capacity contribution of rate-adaptive relaying combined with frequency recycling.

3.6 Load Model

Mainly two types of load models are considered in this thesis:

1. A constant bit rate (CBR) load and every terminal is assumed to have an have a CBR traffic from an infinite source
2. A variable bit rate (VBR) load with every terminal is assumed to have a randomly sized traffic such as email, SMS, HTTP or video clips to send.

The communication link is assumed to be symmetrical and hence the treatment of up- and downlink traffics is similar.

3.7 System Metrics

In a cellular network with terminals communicating either to the AP or with each other, the system resources such as frequency spectrum and time are utilized in various ways. Here, the time resource is shared among the communicating terminals either equally, independent of their location from the AP, or is allotted according to the distance they are far from the AP. In the end, the resource utilization is evaluated by the system metric *capacity per available bandwidth*. Based on two fairness schemes to be discussed in Section 6.4, the capacity per available bandwidth in a cellular network is defined as follows:

Definition 3.2 *Consider a cell with N terminals uniformly distributed inside the cell and ready to send their traffic to the AP. Assume that the terminals share the TDMA frame equally and send their traffic with the maximum possible data rate,R, at some maximum packet error rate PER. The capacity of the cell at the AP per available bandwidth is defined as the joint traffic that is passing through the AP given by:*

$$C = \max_{(R_1,\ldots,R_N)} \sum_{i=0}^{N} (1 - \mathrm{PER_i}) R_i \tag{3.12}$$

Definition 3.3 *Consider a cell with N terminals uniformly distributed inside the cell and are ready to send their traffic to the AP. Assume that all the terminals are allowed to send uniform amount traffic* *in the TDMA frame with the maximum possible data rate,* R*, at some target packet error rate* PER*. The capacity of the cell at the AP per available bandwidth is defined as the joint traffic that is passing through the AP given by:*

$$C = \sum_{i=0}^{N} \left(\frac{1 - \mathrm{PER_i}}{\max_{(R_1,\ldots,R_N)} \sum_{i=0}^{N} \frac{1}{R_i}} \right) \tag{3.13}$$

Definition 3.4 *Consider a cell with* N *terminals uniformly distributed inside the cell and have a variable bit rate load with a randomly sized traffic to send to the AP. Assume that the terminals share half of the TDMA frame according to their traffic size and the other half of the frame equally. Let each terminal sends a traffic size of* x_i *at the maximum possible data rate* R_i *and packet error rate* $\mathrm{PER_i}$*. The capacity of the cell at the AP per available bandwidth is defined as the joint traffic that is passing through the AP given by:*

$$C = \max_{(R_1,\ldots,R_N)} \sum_{i=0}^{N} \frac{1}{2} \left(\frac{1}{N} + \frac{x_i}{\sum_{i=0}^{N} x_i} \right) \cdot (1 - \mathrm{PER_i}) R_i \tag{3.14}$$

In other words the capacity of a cellular network is defined as the amount of data correctly received or sent by the access point for a given communication time. In terms of users, the capacity of a cell can also be defined as the number of mobile terminals that communicate with the access point at a given communication for a given minimum throughput. Target packet error rate is the maximum allowable packet error rate set in the system. For e.g, in HiperLAN/2 the maximum allowable PER is 10% [29] and there will not be any communication if the actual PER between two communicating terminals is larger than the target PER. The upper and the lower bound of the achievable capacity per available bandwidth at the AP is as follows:

1. Upper capacity limit

   An upper capacity limit is attained when a single terminal is allowed to send over the entire TDMA frame with the maximum possible data rate which is 54 MBit/s. Allocating the entire frame to a single terminal only to achieve the upper capacity is not a fair way of sharing the available resource, especially when there are other transmitting terminals within the cell. Hence the upper capacity limit needs to incorporate all the transmitting terminals sharing the available resource according to fairness criteria. Consequently, if every transmitting terminal is allowed to share the TDMA frame slot uniformly and is capable of communicating with the AP at a rate R_{max} with a negligible PER, the upper capacity limit at the access point is:

$$C_{upper} = \sum_{i=0}^{N} R_{max_i} \tag{3.15}$$

And if every transmitting terminal is allowed to send a maximum uniform traffic size ξ_{max} with a negligible PER, the upper capacity limit at the access point is:

$$C_{upper} = \sum_{i=0}^{N} \xi_{max} = \xi_{max} \cdot N \tag{3.16}$$

2. Lower limit capacity

   It is trivial that the lower limit capacity at the access point is zero if the PER is too high to sustain the communication within the cell. But assuming that there is communication at the maximum allowable PER, the lower capacity limit is set forth as:

$$C_{lower} = \sum_{i=0}^{N} (1 - \mathrm{PER}_{\mathrm{max_i}}) R_{min_i} \tag{3.17}$$

   if all terminals are allowed to share the TDMA frame slot uniformly and if they are communicating with the access point at the slowest rate R_{min} with maximum allowable target packet error rate $\mathrm{PER}_{\mathrm{max_i}}$. Similarly, if all transmitting terminals are sending a non-zero minimum uniform traffic size ξ_{min} at $\mathrm{PER}_{\mathrm{max_i}}$, the lower capacity limit is set forth as:

$$C_{lower} = \sum_{i=0}^{N} (1 - \mathrm{PER}_{\mathrm{max_i}}) \xi_{min} \tag{3.18}$$

Here, $C_{lower} < C \leq C_{upper}$ and usually C is much less than the upper limit. Also, when the packet overhead is considered, the lower limit capacity becomes even smaller than the expression indicated.

Chapter 4

Related Work

4.1 Introduction

There have been several studies that have contributed to the understanding of network capacity and utilization in multi-hop wireless networks. Among these, the ones which are closely related to this thesis are presented in this chapter. The remainder of this chapter is organized as follows. In Section 4.2, the theoretical approaches used to analyze the capacity of wireless networks are described. Section 4.2.3 discusses the rate adaptive and multi-channel approaches used in wireless networks. Fairness and scheduling techniques are described in Section 4.4. Improving throughput by means of controlling the power is presented in Section 4.3. In the end the summary of the different related works is presented and an outlook of the succeeding chapter is given.

4.2 Information Theory Approach

4.2.1 Ad Hoc Networks

The capacity of wireless ad hoc networks has been fundamentally studied by GUPTA and KUMAR [41, 42]. They considered an ad hoc wireless network model where n identical nodes on a disk of unit area communicate over the wireless channel and cooperate to relay traffic. The physical model they assumed is one where each link operates at a fixed data rate utilizing a finite bandwidth and power. They also defined a protocol and physical model that guarantees successful simultaneous transmissions in the same channel. The key result under these assumptions is that the maximum achievable per node throughput is $O(\frac{1}{\sqrt{n}})$ and as the number of nodes n per unit area increases, the per node throughput goes to zero. The result remains the same whether nodes transmits in a given common channel or in a splitted sub-channel. Besides, they studied the case when relay nodes are deployed randomly in the network with no traffic of their own but used purely for relaying purpose.

Their result showed that to gain an appreciable increase in capacity, the number of relay nodes to be deployed should be very large. For e.g., for $n = 100$ active nodes, at least 4476 relay nodes are needed to have a five times increase in throughput. If the per node throughput of ad hoc networks goes to zero as the network size grows, it is inevitable to ask whether large ad hoc networks are feasible at all. A simulation based study in [69] examined the feasibility of large ad hoc networks by combining different traffic patterns and network configurations such as simple chain and lattice structures. The result achieved pointed out that if the average communication distance grows with the network size, the per node capacity rapidly decreases and the network capacity is marked below optimal. Nevertheless, if nodes send only to nodes within a fixed radius, independent of the network size, the per node capacity is likely to scale to larger networks. This shifts the question of scalability to likely locality of communication in the network.

These pessimistic results has initiated several works that have presented different results under different assumption. Our motivation is also to show the possibility of increasing the aggregate capacity of the wireless network even if the per-node capacity decreases as the number of nodes in the cell increases. This is done by jointly considering transmission rate, transmission power, routing and scheduling in a relaying network, which are traditionally treated independently as separate elements of the physical, link and routing layers.

The results achieved in [41] are under the assumption of point-to-point communication, i.e., during any given time slot, one node transmits to exactly one other node and the latter considers all other incoming signals purely as noise. Using similar physical models of wireless network but applying arbitrarily complex network coding such as multi-access and broadcast codes as opposed to point-to-point coding GASTPAR and VETTERLI [36] showed that the capacity of wireless networks with n nodes under a relay traffic pattern behaves like $\log n$. Essentially the achievable rate is constant independent of the number of nodes.

Similarly, GROSSGLAUSER and TSE [40] introduced mobility to the model used in [41] and evaluated the achievable per node throughput. Their main result shows that the average long term per node throughput can be kept constant even as the number of nodes per unit area increase. In other words, there is no significant loss in per node throughput in terms of growth rate as a function of n, number of nodes. But the caveat of this result is that the attained long term throughput is averaged over the time scale of node mobility and hence delays of that order will be incurred. BANSAL and LIU [8] tried to compromise between [42] and [40] by proposing a routing algorithm that helps to scale the per node throughput while at the same time bounds the delay incurred in delivering the relayed traffic. An Optimal throughput-delay tradeoff is also presented for both static and mobile networks in [35, 94].

For arbitrary number of nodes and topology in wireless ad hoc networks, capacity regions are defined in [111, 113]. These regions describe a set of achievable rate combinations between all source and destination pairs in the network under various transmission strategies, such as variable rate transmission, single- or multi-hop routing, power control and successive interference cancellation. As a figure of merit, they defined the uniform capacity C_u of a network under a given transmission protocol as the maximum aggregate communication rate when all nodes wish to communicate with all other nodes at a common rate. Their numerical result indicated that multi-hop routing without spatial reuse increased the uniform capacity C_u of a single-hop routing by 242% and when spatial reuse is introduced to the multi-hop routing, the C_u is further increased by 26%. On the other hand, the gains from power control are significant only when variable rate transmission is not used. Although it cannot be directly compared due to the fact that their approach is entirely on a purely ad hoc network, their main result has good coherence with the results in this thesis in a way that relaying gives a better capacity to the network and if frequency reuse is applied to the relaying, the network capacity can further be improved.

As opposed to the typical communication model assumed in wireless ad hoc networks, where packets are sent in a pipelined manner and several packets belonging to an end-to-end flow simultaneously wait to be served at different stages along the flow's path, a non-pipelined relaying strategy is used in [118] to improve the throughput performance of the wireless ad hoc network. The authors argued that through non-pipelined relaying, it is possible to reduce the degree of contention and impact of route failures which amount to the improved network utilization. An increase per node throughput with increase in node density is presented in [87, 130] from another perspective. The results are achieved by use of ultra wide band links which have properties, such as extremely large bandwidth and low power, that makes them different from the existing wireless technologies.

4.2.2 Hybrid Networks

Hybrid wireless networks are sparse networks of base stations in an ad hoc network. The stations are assumed to be connected by a high bandwidth wired network and act as relays for wireless nodes. They are neither data sources nor receivers [73]. These hybrid networks present a tradeoff between traditional cellular networks and pure ad hoc networks in that data may be forwarded in a multi-hop fashion or through the infrastructure [53].

Akin to pure ad hoc networks, the scalability of hybrid networks is studied in [65, 73]. LIU et al. [73] identified two different routing strategies and obtained analytical expressions for the capacity of hybrid networks. They assumed m base stations and n nodes, each capable of transmitting at W bit/s over the wireless channel. In their first routing type, a node sends data through the infrastructure if the destination node is outside of the cell where the source is located, otherwise the data are forwarded in a multi-hop fashion as in an ad hoc network. Under this strategy, if m grows asymptotically slower than $\sqrt{n}$, the maximum throughput capacity is $O(\sqrt{\frac{n}{\log\frac{n}{m^2}}}W)$ and the benefit of adding base stations is insignificant. However, if base station is added faster than $\sqrt{n}$, the maximum capacity is $O(mW)$, which increases linearly with the number of base stations. Similar results are obtained for the second routing strategy, where a node uses infrastructure to send data according to some probability. Here, if m grows slower than $\sqrt{\frac{n}{\log n}}$, the maximum throughput capacity has the same asymptotic behavior as in a pure ad hoc network. In short, if the number of base stations scales slower than some threshold, the throughput capacity is dominated by the contribution of ad hoc transmission, in which case the benefit of adding base stations is minimal. If the number of base stations scales faster than the threshold, the capacity is dominated by that of infrastructure. Thus, the scalability of the capacity of hybrid networks lie at the cost of high investment in infrastructure, which still needs a tradeoff.

KOZAT and TASSIULAS [65] also obtained an expression for transport capacity of hybrid networks. They assumed that the number of ad hoc nodes per access point is bounded above. Likewise, each wireless node, including the AP, is assumed to be able to transmit at W bit/s using a fixed transmission range and n ad hoc nodes, excluding the AP, constitute a connected topology graph. They obtained a per node throughput capacity of $O(\frac{W}{\log n})$ which is higher than that of a purely ad hoc network in [41]. A similar study in [2] showed how the hybrid network capacity scales better with power control. The result indicated the possibility to provide a static throughput capacity of $O(\frac{1}{logn})$ for each node and one can guarantee a throughput of $O(1)$ for some $O(n)$ nodes in the network. Here the nodes are allowed to choose their power levels in contrast to the hybrid network shown in [65]. The authors in [65] argued that the infrastructure based approach provides a simpler mechanism that has more practical aspect than obtaining a better capacity figures by complex network coding or exploiting mobility in the network. A similar result is obtained in [110]. The multi-hop cellular wireless network considered in this thesis is in a way a hybrid network since it comprises of both ad hoc networks and base stations. But the nodes are more bind to the access point than having an ad hoc property as in the hybrid network, i.e., nodes within a cell should always communicate with their AP even if they are in an ad hoc relaying mode.

4.2.3 Multi-Hop, Multi-channel and Multi-rate Approaches

The notion of capacity has also been studied in several contexts for both TDMA and CDMA based wireless networks. Various approaches have been proposed to improve the network capacity. Some of these approaches are multi-hopping, transmission rate control, channel reuse, power control, spatial multiplexing (e.g, [46]), resource optimization (e.g., [12]) or the combination two or more of these using queueing analysis (e.g., [4]), etc. In this section, multi-hop, multi-channel and multi-rate approaches are discussed.

Multi-hop Architectures

A wireless network architecture called the multi-hop cellular network (MCN) is proposed in [72]. It combines the features of conventional cellular and ad hoc network architectures. The merits of this architecture is multifold: the transmission range of communication can be reduced, terminals can communicate without involvement of base stations, multiple packets can be transmitted within the cell and paths are less vulnerable as the base station can help reducing the wireless hop count. The performance of the architecture is studied by applying the DCF access scheme of the IEEE 802.11 MAC protocol. The analysis and simulation result showed that the proposed architecture had a superior performance than the conventional cellular network architecture. The increasing order of mean number of channels for simultaneous transmissions and mean hop count is illustrated as the transmission range decreases. The downside of the approach is the need for packets to be send multiple times to arrive at destinations that consume more bandwidth which eventually limits the system throughput. The need to find an appropriate operational value of mean channel and hop count for optimal system performance is still left open as it is not easy to increase number of transmission channel as needed. The issue of transmission range minimization with respect to higher transmission rate is not explicitly addressed in the study. Conceptually, the use of both cellular and ad hoc features to improve the performance of the conventional cellular network is similar to the work done in this thesis though the approach remains different.

In a similar manner, QIAO and WU [95] addressed the issue of increasing system throughput by introducing ad hoc relay stations in their integrated cellular and ad hoc relay (iCAR) architecture. These ad hoc relays are placed at strategic locations to divert traffic from heavily congested cells to other less congested cells [123]. Such special devices are cost ineffective as they are fully utilized at times of heavy congestion only. In comparison, the relaying nodes in this thesis are the terminals themselves with their own traffic to send to the AP and hence no additional device is required for the system. Issues related to cellular and ad hoc architecture are also studied in [78, 122].

Multi-hop and Multi-channel Networks

The performance comparison of cellular and multi-hop wireless networks in terms of network capacity, end-to-end delay, power consumption, per node fairness and the impact of mobility is presented in [54]. The simulation result showed that the ad hoc multi-hop network performs better in terms of throughput, delay and power. On the other hand, they are unfair to network nodes and the performance of the network suffers in the event of mobility due to the increase of dropped packets and route rediscovery overhead . The authors suggested a hybrid network which takes advantage of both ad hoc and cellular network for a better throughput and fairness. However since their work is limited to single cell, they did not consider interference in the model, which is one of the factors that impede capacity. A similar performance investigation was also done on multi-hop wireless networks in [44].

In the context of IEEE 802.11 LI et al. [69] have performed a simulation study of the relaying behavior of 802.11's distributed coordination function (DCF). They considered different scenarios (regular and random placement of nodes) and attempted to communicate packets over multiple hops between different nodes. Their main result illustrated that although IEEE 802.11 DCF finds reasonably efficient communication schedules even in complicated networks, only local communication patterns can be efficiently supported (in agreement with the theoretical results derived by [41]). The per node throughput in [41] indicated that there is no change in result whether terminals are transmitting over a given channel or over multiple sub-channels. However, the results in [68] showed the dependency of capacity on the ratio of the number of channels to the number of interfaces per node and not on the exact number of any one of them. Other simulation results also showed that the use of multi-channel gives additional improvement to the multi-hop networks as compared to single-channel network. (e.g., [57, 70, 71, 103, 108, 115, 116, 125]). Whereas these improvement results are valid for networks with CSMA-based medium access control protocol, the results are not directly comparable to network model described in this thesis.

RAPP [97] used the HiperLAN/2 system to demonstrate how its throughput and QoS can be increased. The author exploited the silent periods in the TDMA transmission frame by placing them asymmetrically in the entire frame; as a result the interference situation in certain areas of the frame can be improved for both up- and downlink transmissions in co-channel radio cells. The TDMA based system and assumption used to study the performance improvement is somewhat similar to the one in this thesis, however, the approach is entirely different. In [26] a concept of forwarding mobile terminal is introduced to add a multi-hopping feature to the HiperLAN/2 network and improve its throughput and coverage. This forwarding terminal is a new element between the AP and the remote terminal which is recognized as an AP by the remote terminals and as a terminal by the AP. However, the new element is different from the other terminals and requires additional functionality to implement the forwarding.

Relaying as a means to provision a better network performance is studied for CDMA networks as well in [13, 47]. The authors showed how the overall performance of a CDMA relay system depends on the node density and the relative load. They suggested that direct transmission eventually becomes favorable when contemplating capacity improvement. In fact, our results also suggested that relaying only may not be enough to bring about a significant increase, it should be substantiated with appropriate scheduling and multi-channel transmissions in a cell.

4.3 Power Control

As indicted earlier one factor which affects the performance of a radio network is interference. This factor is greatly influenced by the transmission power terminals are using in the network. Applying a power control mechanism may have a dual advantage in mitigating interference and in prolonging the battery life of the network device. The following section presents the power control studies done with respect to capacity improvement.

The need for computing optimal transmission ranges in radio networks is indicated in [52, 63]. These earlier works illustrated that the use of large transmission radii has a high degree of connectivity, but there will be a corresponding loss of channel throughput due to interference. Conversely, use of shorter transmission radius increases the channel throughput but there is a corresponding increase in number of hops for the traffic to reach its destination. Large number of hops tend to reduce the effective channel throughput owing to increased internal traffic. The authors suggested transmission strategies which use transmission power control to achieve a better performance. KLEINROCK and SILVESTER [63] also showed that the optimal transmission radius should be proportional to $\sqrt{n}$ – a cornerstone of the seminal work of GUPTA and KUMAR [41].

The benefit of power control to improve capacity in multi-hop wireless networks has been studied both theoretically [41] and with simulations (e.g., [86]). The idea used in [86] is to reduce the tranmission range and transmission power so that more intermediate hops are allowed in the network, which again results in an increased number of simultaneous transmissions. A basic pathloss model which causes a signal to attenuate with distance on the order of $\frac{1}{d^{\alpha}}$, is used to examine the degree of transmission power reduction (where d and α are the communication distance and the pathloss factor, respectively). The results confirmed the benefits of power control mechanism for a better capacity, but the approaches are limited to a single channel IEEE 802.11 MAC protocol. The use of multi-channel and multi-rate transmissions are not explicitly considered.

The significance of power control has also been studied in combination with network optimization parameters such as transmission rate, channel-to-interference ratio (C/I), routing and link scheduling. In [62], a linear optimization model is proposed to solve the combined power control and transmission rate selection in cellular networks. The authors showed performance improvement in terms of throughput, outage probability and transmission power consumption for CDMA system. However, the results are for direct mobile to base station communication only and the problem of relaying is not addressed. In [92] a heuristic joint resource allocation algorithm (JRAA) is presented to solve the channel, base station and power assignment in wireless networks. Unlike cellular networks whose base stations are preassigned, an arbitrary number of base stations and mobiles are assumed in the network. They identified all possible subsets of mobiles that can make use of the same channel and then assigned minimum number of channels, base stations and corresponding powers. Their numerical result fortified that the joint resource and power assignment brings a considerable system capacity improvement. A similar resource allocation scheme with power control is presented in [5]. But again the results are applicable to the traditional cellular networks and do not extend to multi-hop cellular networks.

SHEPARD [101] showed a decentralized channel access scheme for a scalable packet radio network that is free of packet loss. In this scheme, a power control algorithm is used to reduce the excessive power used for transmitting to stations which are closer to the maximum range, and to transmit with the same average power density as before, keeping average signal to noise ratio constant. With reduced power, as long as the low power level can still deliver a sufficient signal to noise ratio at the receiving station, interference to other stations can be reduced and the signal to noise ratio in receivers at the other station can also be increased. Minimum-energy routing is also used in their channel access scheme to route packets traveling more than a distance of $2\rho^{-1/2}$, where ρ is average distribution density of stations, through intermediate stations to minimize packet loss. One of the criteria used to determine routes is choosing the nearest hop, which produces less interference. Basically, one of the relay selection criteria used in this thesis is the nearest hop, but it is rather a throughput-aware routing than energy. Issues related to power control in IEEE 802.11 have also been studied in [85, 102, 114].

Likewise, power control in HiperLAN/2 networks is studied in [66]. The author argued that the use of maximum transmission power usually results in a data rate that is higher than the required one.

This makes the MAC frame to be filled only partly. A combination of link adaptation and power control can be used to distribute and reduce the unused parts of the MAC frame and transmission power is increased only when more capacity is required. The approach targets MAC-Frames which are not completely filled and is effective for the case of time variant loads.

4.4 Fairness and Scheduling Approach

In wireless systems, efficient resource management plays an important role in meeting stringent quality of service and high system performance requirements. For this purpose, a good understanding of user requirements and subsequently good resource allocation techniques are needed. This section discussed issues related to resource allocation and fair scheduling of resources.

LIU et al. [74] used fair resource allocation mechanism to improve the efficiency of wireless network utilization. Their argument is that it is possible to increase the overall throughput of the network by allowing terminals close to the base station to transmit with high transmission rate and power, being unfair to other terminals, which should not be the case. But instead, they come up with a transmission scheduling policy that exploits the time-varying channel conditions and maximizes the system performance stochastically under a certain resource allocation fairness constraint. As a result, they obtained 20% to 150% performance improvement compared to that of round-robin policy.

A joint routing, link scheduling and power control scheme is proposed in [19] to support high data rates for broadband wireless multi-hop networks. The authors tried to find subsets of simultaneously active links and the associated transmission powers to minimize the total average transmission power expended across the network. They regarded minimum data rate per link and peak transmission power per node as constraints so as to achieve target capacity on links. Although their scheme achieved a higher throughput than other radio resource allocation approaches, the energy efficiency is low. Likewise KODIALAM and NANDAGOPAL [64] used a joint routing and transmission scheduling approach to characterizing the rates that are achievable in wireless networks. An interference model similar to IEEE 802.11's RTS-CTS scheme is used to silence the neighborhood of the transmitter and receiver. Linear optimization and edge-graph coloring techniques are used to solve the routing and scheduling problem. For nodes that transmit to or receive from at most one node at a time, the authors suggested algorithms which are within 67% of the optimal solution in the worst case.

Fairness has also been important issue in scheduling the wireless network resources [6, 7, 55, 75, 77, 96, 109, 117]. In most of the papers, different fair scheduling mechanisms are proposed to allocate bandwidth in proportion to the weights of packet flows sharing the channel. In [53] a MAC protocol is proposed for IEEE 802.11 which supports prioritized per node fairness to improve throughput and fairness. They argued that although the per node fairness model adopted for cellular networks is apt for the environment, it is not suitable for ad hoc networks where nodes cooperatively act as relays for flows belonging to other nodes in the network. So they presented a load balance routing algorithm to reduce the average degree of flow multiplexing on a single link by taking capacity and hop count into account along the path. A fair resource sharing mechanism is proposed in [80] for inter-working and co-existence of IEEE 802.11a and HiperLAN/2 multi-hop networks. The motivation is both networks operate in the unlicensed 5 GHz band, consequently, data throughput, QoS levels and connectivity will be degraded and unpredictable if the networks are not able to coordinate the competing access to radio resources. Thus, the authors used an integrated protocol which is capable of operating in both 802.11a and HiperLAN/2 modes and is capable of centrally coordinating the communication. They also used the QoS enhancement in IEEE 802.11e to maintain the QoS support for both systems. In this thesis, fairness has been an underlining issue as well. It reflects two perspectives on how to share system resources – provider oriented or customer oriented, as discussed in Section 6.5.

4.5 Routing, Interference and System Performance

This section discusses issues related to relay path selection scheme and interference based system performance.

The work done by SRENG et al. [105], [104] is relevant to this thesis. The authors use a simulation approach to investigate the impact of different relay node/path selection and channel selection schemes (from among channels already used in adjacent cells) for relaying purpose on the users coverage, with and without power control. In their simulation, they used a TDMA cellular network with a central controller and mobile nodes having a down-link communication. The mobile nodes are allowed to have two-hop relaying in order to improve coverage in the cellular network. They proposed a path selection scheme based on geographic distance and pathloss and a channel selection scheme by computing the carrier-to-interference ratio of every link. In all of their simulation, they assumed fully loaded cells with continuous traffic and the number of available channels per cell increases linearly with the number of subscribers. Moreover they considered larger rectangular cells ranging from 2 × 2km to 400×400m. They implemented their proposed schemes for both Noise-limited and interference-limited environments, with and without power control and with and without adaptive modulation. Their result shows a considerable improvement in coverage. As a side remark, they also observed the impact of adaptive modulation and power on average node throughput due to their relay/ path and channel selection algorithms.

The concept of using adjacent channels for relaying purpose is somewhat similar to this thesis. Our frequency recycling scheme takes the scarcity of communication bandwidth into consideration unlike in their case which assumes available channel per cell that increases linearly with the number of nodes. Power control in their sense is step by step increasing of the transmission power of the terminals until a certain maximum allowable level. In our case, the transmission power is always the minimum power needed to transmit the traffic for the selected data rate and specified target packet error rate.

The size of the cell we are using is mainly dependent on the maximum transmission radius that a terminal can communicate to the AP with the lowest modulation rate. So every terminal in each cell is always reachable by the AP. Whereas in their work, they use continously increasing cell size, in which terminals may or may not be reachable by the base station. So we both have a different target: optimizing capacity in our case and optimizing coverage in their case.

Adaptive modulation in combination with scheduling of the TDMA frame for concurrent transmission by at the same time maintaining fairness among mobile terminals for capacity optimization is uniquely identified in this thesis. Although they assess the characteristics of average node throughput per cell for adaptive modulation and power control, the model they are taking is different from ours and it is difficult to draw comparisons. But they at least showed an improvement in average node throughput when relaying is used–which confirms to the results that we also have.

A similar performance study is done in [107] for TDMA-based fixed cellular networks with inadequate single-hop coverage. Relaying in multi-frequency channel is used to obtain a significant coverage boost. The transmitting terminals themselves act as a relay terminal and no separate terminal and relaying channel is dedicated in the system. The result shows coverage enhancement as the number of available channels increases.

The impact of interference in multi-hop wireless system is studied in [22, 45, 56, 107]. In [56] the interference is modeled as a conflict graph to define the upper and lower bounds on the optimal network throughput. Their optimal routing computation outperform that of shortest path routing for any given network configuration and load as it is based on less interference prone path. However, the results are under an unrealistic assumption that there is an entity with full knowledge of each packet transmission and capable of optimally scheduling every transmission.

4.6 Summary and Outlook

In the earlier chapters, an overview of the existing wireless technologies, network components and system performance are presented. It has also been mentioned that the demand for wireless networks is enormously growing and it is the current challenging issue to satisfy the requirement of the user. One of the most important requirements of wireless networks is the provision of high system capacity. Methods of increasing the system capacity is, thus, vital area of research. This chapter is essentially dedicated to discussing the significant researches done to improve the capacity of wireless networks.

The main ideas of the present chapter are recapitulated as follows: The capacity of ad hoc wireless network is profoundly studied in [41, 42]. Their result is somewhat pessimistic as the per node throughput for very large nodes actually goes to zero even with perfect scheduling. From information theory point of view works of [36, 112, 113] suggested the possibility of using relaying in the ad hoc network to improve the fundamental result achieved in [41]. The assertion in this thesis is also to show the possibility of increasing the aggregate network throughput by making use of the increase in nodes as a means to find an optimal relaying node for nodes that far from their AP.

In [2, 53, 65, 73] hybrid wireless networks are studied to make use of the features of pure ad hoc and the infrastructure-based cellular network. The results suggested the possibility of hybrid networks to provide a minimum per node throughput which is greater than zero for a very dense network. Parallel to this, in [72, 78, 123, 124] a multi-hop cellular network architecture is studied to adopt the relaying capability of nodes in ad hoc network to the traditional cellular network to improve the network throughput. The relaying nodes as in [123], for example, are not the transmitting terminals themselves but special devices placed at strategic locations to divert traffic from heavily congested cells to other less congested cell. Such relay nodes are fully utilized when the network is fully loaded and they are cost inefficient. The motivation of this thesis is also the notion of adopting the relaying features of ad hoc to the traditional infrastructure-based cellular network. But it is not limited to allowing node in ad hoc networks to use the infrastructure as in hybrid networks or use relay nodes in the network only to assist in relaying. Rather, the access point is the central controller of every communication in the cell it belongs and relay nodes are the transmitting nodes themselves which assist in relaying whenever it is needed.

The structure of the succeeding chapters is as follows: Chapter 5 presents an analytical model to estimate the capacity that can be achieved by introducing relaying in wireless cellular network. To get an intuition on how the system performance is influenced by factors such as transmission power, noise and interference, the network is separately studied as transmission power limited and interference limited system. In the transmission power limited system case, where there is only one AP and a number of terminals in a cell, terminals are allowed to send their traffic either directly or via a relay terminal using a limited transmission power. Here, the noise and the limited transmission power are predominately affecting the throughput of the network under the assumption that there is no or negligible interference. In the latter case, two APs which are using the same frequency are used to study the effect of interference in the system throughput. Relaying is also introduced to see by how much the interference problem can be ameliorated to bring about capacity improvement. Though simplified the analytical models may seem, they have keen significance to hypothetically see if relaying is wholesome and to what range it should be applied. After analyzing the potential benefit of relaying, the natural question that should follow is how often relaying terminals are available which are ready to assist far terminals. Thus, the estimation of the probability of finding such relaying terminals in a network is presented at the end of this chapter.

Once the potential benefits of relaying are identified, the separately studied analytical models need to be extended to a realistic network model. There are various factors which characterize the realistic wireless network, some of which are described in Chapter 3. Analytically solving the network model by combining all these factors is rather cumbersome. Therefore, simulation-based solution is opted to study the wireless network. In Chapter 6, the novel approaches used to improve the capacity of the cellular wireless network through relaying are described. These include the transmission power and rate adaptation problem for each individual link, the routing problem to select a relaying terminal for each terminal or to decide not to relay at all and the problem of fairly scheduling the system resource. In effect, it is an integrated approach to solve the physical, the link and the network layer problem. Moreover, appropriate frequency recycling schemes are described in the chapter. A numerical example is also included to verify the possible outcomes expected from these approaches.

Following this, the simulation environment and the corresponding simulation results are presented in Chapter 7 and 8. In the end, a comprehensive discussion is given on the implications of the simulation results and their relationship with that of the analytical model.

The approaches proposed to improve the network capacity are valid for a static topology assuming that terminals are stationary for most of the time. If the terminals in the network are not stationary anymore, the proposed mechanisms may not yield an optimal system performance. To maintain an optimal result, there should be a frequent update of the topology and routing information. Such issues related to mobility are presented in Chapter 9. Furthermore, a mathematical analysis is presented to see if the wireless network model considered can make use of mobility to get a additional capacity boost.

Apart from system capacity, energy efficiency is also a critical issue in wireless networks. Relaying makes relaying terminals busy in sending both their own and relayed traffic. This is, in fact, done at the cost of relaying terminal’s battery capacity. Methods of improving capacity may have a diametric effect on the energy efficiency. Chapter 10 takes a look at the effects of the proposed mechanism in the total power consumption of the network. Finally, Chapter 11 sums up the whole discussion and points out open issues and future works.

Chapter 5

Analytical Estimation of Cell Capacity

5.1 Introduction

It has been pointed out in the earlier discussions that relaying traffic as opposed to direct transmission has prominent advantages. Primarily, interference is mitigated as terminals close to the maximum range can now transmit their traffic to the access point via relaying terminals which are closer to the access point with reduced transmission power. Secondly, for terminals that are already using maximum transmission power, at least the packet error rate can be reduced. These effects should promise an increase in capacity–the total amount of traffic which can be supported by a set of terminals in a given amount of time.

In order to have a first insight to what extent relaying can promise an improvement, relaying is introduced into the basic capabilities of the wireless network and the system is separately studied as *transmission power limited* and *interference limited system*. In transmission power limited system, the cellular network is assumed to have a very large frequency reuse distance and the effect of interference from simultaneous transmissions in co-channel cells is almost negligible. The main factors that affect the performance of this system are the strength of the transmission power and noise. Studying such a system makes it easier to see the relation between transmission power and relaying.

Likewise, in interference limited systems, interference is the dominant factor that affects the performance of the network and the effect of noise can be ignored. By analyzing such a system, it is possible to see how relaying can influence the effect of interference by reducing the communication distance.

In this chapter an analytical model is defined and the potential benefit of relaying is mathematically analyzed for both systems. In addition, an analytic model is presented to estimate the possibility of finding a relaying terminal with in the cellular network which can potentially yield the presumed advantage.

5.2 Transmission Power Limited System

In this section a transmission power limited system is studied by taking a single cell with an AP. Terminals in the cell are allowed to send their traffic at a time, either directly or via a relaying terminal. The effect of interference is assumed to be negligible. The extent to which the terminals can vary their transmission rate and power to achieve a higher throughput is the concern of this section. The remainder of the section is organized as follows: In Section 5.2.1 the analytic network model is described and the mathematical analysis is given in Section 5.2.2 to compare the total throughput obtained for both direct and relay communication. Finally, Section 5.2.3 presents the numerical result.

5.2.1 Model

Consider the cell model shown in Figure 5.1 to investigate the capacity improvement of relaying transmission over the direct transmission case. The cell contains an access point at the center of the cell, a relaying terminal MT_1 at a distance d close to the access point and three other transmitting stations d' away from the relaying terminal. All the terminals are assumed to have their own traffic coming from infinite queue source. For simplification, we assume that the three "far" terminals MT_2 to MT_4 are located very close to each other, so that the differences in distance to MT_1 or AP are negligible. This scenario is chosen to easily vary the different transmission distances with respect to the AP.

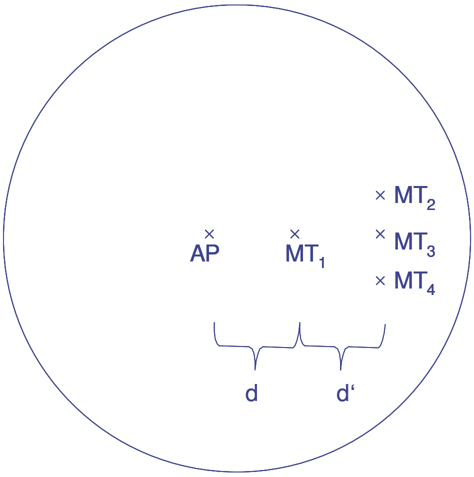


Figure 5.1: System scenario

Let a transmission power P_T be used to transmit the traffic load from the terminals either to the access point or to the relaying terminal. The pathloss model shown in Equation 3.6, which depends on d, the distance between sender and receiver terminal, and the pathloss coefficient α, is used to relate the transmission power P_T to the received power P_R:

$$P_R = \frac{P_T}{10^{\frac{1}{10}L}} = \frac{P_T}{10^{\frac{1}{10}(46+10\cdot\alpha\log 10(d))}} \tag{5.1}$$

$$P_{R[dB]} = P_{T[dB]} - (46 + 10\cdot\alpha\log 10(d)) \tag{5.2}$$

Moreover, the transmission media is shared in TDMA manner and no two terminals transmit at the same time using the same frequency band. Hence, there is no interference within the cell and the only sources of packet errors are noise and limited transmission power. The signal to noise ratio is expressed simply as $\mathrm{SNR} = P_R/N$ or, if SNR is expressed in dB, $\mathrm{SNR} = 10\log_{10}(P_R/N)$.

The analysis is carried out based on the interference-to-error-rate mapping data for HiperLAN/2 [61]. An exponential curve fitting [91] is used to mathematically describe the mapping and the PER is:

$$\mathrm{PER}(s,i) = 10^{a_i s^2 + b_i s + c_i} \tag{5.3}$$

where $s = 10\log_{10}\mathrm{SNR}$, the SNR in dB, i is the index of modulation used in HiperLAN/2 and a_i, b_i, and c_i are as defined in Table 5.1, which also shows name and nominal bit rate (NBR) for these modulations. An appropriate cutoff is chosen to ensure that the PER remains between 0 and 1.

Index i	Modulation (code rate)	NBR $(\mathrm{MBit/s})$	a_i	b_i	c_i
1	BPSK (1/2)	6	-0.00826140426805	-0.06376668709407	-0.19668486235428
2	BPSK (3/4)	9	-0.00691007462078	-0.01170647235394	-0.10819588784435
3	QPSK (1/2)	12	-0.00961459243554	+0.00515845543333	-0.14211581761582
4	QPSK (3/4)	18	-0.00689529575429	+0.02765247907588	-0.10966369423267
5	16QAM (9/16)	27	-0.00783459997375	+0.08195372387766	-0.29870707399130
6	16QAM (3/4)	36	-0.00703381983645	+0.11297740455648	-0.53792530792585
7	64QAM (3/4)	54	-0.00623228999252	+0.15392834195885	-0.99488471979605

Table 5.1: Parameters for C/I to PER interpolation

The terminals are assumed to have unlimited data to transmit with the maximum possible data rate to the AP. For simplicity, only the nominal capacity of a given transmission is considered (protocol overhead is not taken into account, which anyway would be almost identical for both direct and relaying). As a perfect transmission over a wireless link is not possible, terminals try to obtain a packet error rate of at most 10% if feasible. Beyond the target error rate, it is assumed that there is not communication.

This traffic assumption is actually conservative as it puts considerable requirements on the relaying node, which has not only to transmit its own traffic but also the other nodes' traffic. This results in the need to use faster modulations between inner terminal and access point (otherwise, queues would overflow in the relay node) with inferior error behavior. Under lower load, this could be avoided, which should in turn improve the performance of relaying.

The performance metric evaluated here is the total amount of traffic, coming from all mobile terminals, that the access point can successfully receive in unit time.

5.2.2 Analysis

In this section, for both direct and relaying communication, the throughput at the access point in the uplink direction is derived. A pre-defined communication schedule is needed to see the effect of direct and relay communication in this TDMA based analysis.

Direct Communication

Each of the four terminals indicated in Figure 5.1 sends its traffic load directly to the AP by selecting one of the seven possible modulations. The corresponding four-time-slot TDMA schedule is shown in Table 5.2.

Slot	Communication
1	$MT_1 \rightarrow AP$
2	$MT_2 \rightarrow AP$
3	$MT_3 \rightarrow AP$
4	$MT_4 \rightarrow AP$

Table 5.2: Direct communication schedule (uplink)

Assuming that the terminals know the distance they are away from the AP, they adjust their transmission power accordingly to keep the PER below some target PER. As the communication distance increases, the transmission power increases as well, but it cannot be increased beyond a maximum limit, for example, a maximum limit of $200\ \mathrm{mW}$ is assumed for this analysis. For distances that require larger transmission power, the PER cannot be maintained and it increases; the throughput decreases accordingly. An example of this situation is shown in Figure 5.2 for two different combinations of α and distance between terminals d. Again, the terminals decrease their modulation rate and readjust their transmission power according to the required SNR until they are within the allowable target PER.

Each mobile terminal adjusts its transmission power according to the following relation in order to maintain the required SNR level

$$\mathrm{SNR}(d) = \frac{P_r(d)}{N} \Leftrightarrow \mathrm{SNR}(d) = \frac{P_t(d)}{10^{\frac{1}{10}(N+46+10\cdot\alpha\log 10(d))}} \Leftrightarrow P_t(d) = \mathrm{SNR}(d)\cdot 10^{\frac{1}{10}(N+46+10\cdot\alpha\log 10(d))} \tag{5.4}$$

where N is noise in dB. As this assumes perfect power control and perfect knowledge of distance between terminals, the SNR and α, it is unrealistic to hope for an actual system to achieve such performance. It only serves as an upper bound to design relaying systems that actually approach the upper bound computed here. This ideal transmission power can only be used as long as it is smaller than $200\ \mathrm{mW}$; otherwise, $P_t(d)$ remains $200\ \mathrm{mW}$ and the resulting SNR and PER should be recalculated.

Based on the given communication schedule, the SNR is computed as:

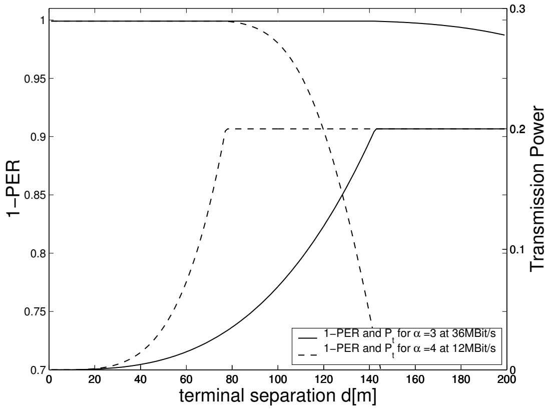


Figure 5.2: Limited Transmission Power and PER as a function of terminal separation for different combinations of α and NBR

Transmitting from MT_1 to AP:

$$\mathrm{SNR}_{MT_1\to AP} = \frac{P_t(d)}{10^{\frac{1}{10}(N+46+10\cdot\alpha\log 10(d))}} \tag{5.5}$$

Transmitting from MT_2 to AP:

$$\mathrm{SNR}_{MT_2\to AP} = \frac{P_t(d+d')}{10^{\frac{1}{10}(N+46+10\cdot\alpha\log 10(d+d'))}} \tag{5.6}$$

The outer terminals (MT_2, MT_3, and MT_4) are assumed to have identical SNR. Once the SNR values of each transmission is computed, the PERs can easily be determined from Equation (5.3) for the employed modulation type j. The throughput $\mathrm{GP}_{\mathrm{direct}}$ for a particular terminal can, then, be approximated as $\mathrm{GP}_{\mathrm{direct}} = (1 - \mathrm{PER})(\text{nominal bit rate})$. This assumes a very simple link layer which does not provide any ARQ protocols. With the assumption that the terminals share the resource fairly, the nominal data rate needs to be divided by four (without considering a downlink phase, which would further scale down the throughput). Thus, the throughput is (using the PER Equation (5.3)):

$$\mathrm{GP}_{\mathrm{direct}}(i, j_i) = (1 - \mathrm{PER}(10\log_{10}(\mathrm{SNR}_{MT_i\to AP}), j_i))\frac{\mathrm{NBR}_{j_i}}{4} \tag{5.7}$$

where i is the terminal under consideration and j_i is the modulation j it uses to send the traffic load to the access point. The total throughput at the access point is finally obtained by adding up the individual terminals' throughput in the cell.

In order to obtain the optimal possible throughput $\mathrm{GP}_{\mathrm{direct}}^{\mathrm{optimal}}$, the best combination of modulations is used. Hence:

$$\mathrm{GP}_{\mathrm{direct}}^{\mathrm{optimal}} = \max_{\substack{j_i \in \{1,\ldots,7\},\\ i=1,\ldots,4}} \sum_{i=1,\ldots,4} \mathrm{GP}_{\mathrm{direct}}(i, j_i) \tag{5.8}$$

where i is the terminal under consideration and j_i is the modulation it uses to send the traffic load to the access point. The total throughput at the access point is finally obtained by adding the individual terminal throughput in the cell.

Relay Communication

The relay communication schedule, shown in Table 5.2, uses the same number of slots as in the direct case. The difference is the outer terminals use the inner terminal as a relay to send their traffic load to the access point. The inner terminal should now transmit both its own and the relayed traffic. Here, it should employ a faster modulation and higher data rate than the overall nominal bit rate of the outer terminals.

Slot	Communication
1	$MT_1 \rightarrow AP$
2	$MT_2 \rightarrow MT_1$
3	$MT_3 \rightarrow MT_1$
4	$MT_4 \rightarrow MT_1$

Table 5.3: Relaying communication schedule (uplink)

This schedule actually limits the flexibility of the relay terminal as it should always use the faster modulation, which is usually error-prone, whether or not the outer terminals have data to be relayed. In the later chapters, we will see how different fairness mechanisms can be used to better optimize the aggregate throughput by choosing relay terminals only when needed.

Similar to the direct case, the transmission power has to be chosen based on the PER requirement. The PER requirement for the two relay hops is slightly different from the overall PER, which remains the same as that of the direct case. To guarantee this, the target $\mathrm{PER}_{\mathrm{relay}}$ of each individual hop in the relaying case should be chosen such that two consecutive transmission have the same target error rate as in the direct case. Thus,

$$(1 - \mathrm{PER}_{\mathrm{relay1}}) \cdot (1 - \mathrm{PER}_{\mathrm{relay2}}) = 1 - \mathrm{PER}_{\mathrm{direct}} \tag{5.9}$$

If we assume similar target error rate at each individual hop, then

$$(1 - \mathrm{PER}_{\mathrm{relay}})^2 = 1 - \mathrm{PER}_{\mathrm{direct}} \Leftrightarrow \mathrm{PER}_{\mathrm{relay}} = 1 - \sqrt{1 - \mathrm{PER}_{\mathrm{direct}}} \tag{5.10}$$

Again based on the schedule, the outer terminals will have SNRs of:

Transmitting from MT_i to AP:

$$\mathrm{SNR}_{MT_i \rightarrow AP} = \frac{P_t(d')}{10^{\frac{1}{10}(N+46+10\cdot\alpha \log 10(d'))}} \tag{5.11}$$

where $i > 1$ is an outer terminal. The inner terminal has the same SNR as in Equation (5.5). Note that the outer terminals are now sending their traffic to a closer relaying terminal and they can maintain the assumed PER level to a considerable distance with the limited transmission power, even though now the target PER for an individual hop is larger than in the direct case.

The determination of throughput is different from that of the direct case as we have to deal with different modulations over the different hops. Consider first the throughput for terminal MT_1, which uses modulation i with nominal bit rate NBR_i to transmit to the access point. Since MT_1 is transmitting only for a quarter of the time, NBR_i has to be divided by four. Moreover, it is only part of the data that belongs to MT_1 in the given time slot– some of the time has to be devoted to the relaying of data received from other terminals. Thus, the actual throughput of MT_1 for transmitting its own data is

$$\begin{aligned} \mathrm{GP}_{\mathrm{relay}}(1, j_1) = & \\ & (1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{MT_1 \rightarrow AP}), j_1)) \\ & \cdot \frac{\mathrm{NBR}_1 - \sum_{i=2}^{4} \mathrm{NBR}_i}{4} \end{aligned} \tag{5.12}$$

where NBR_i is the nominal bit rate of the outer terminal i. The throughput of the outer terminals depends on PERs of both hops (from the terminals to relay and from relay to access point) and is computed as:

$$\begin{aligned} \mathrm{GP}_{\mathrm{relay}}(i, j_i) = & \\ & (1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{MT_1 \rightarrow AP}), j_1)) \\ & \cdot (1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{MT_i \rightarrow MT_1}), j_i)) \frac{\mathrm{NBR}_i}{4} \end{aligned} \tag{5.13}$$

where i is the index of the outer terminal and j_1 and j_i are the modulation indices used by the relaying and outer terminals, respectively. The total throughput at the access point is the overall throughput of each terminal in the cell, the optimal total throughput $\mathrm{GP}_{\mathrm{relay}}^{\mathrm{optimal}}$ is again the throughput achieved with the best combination of modulations, just as in the direct case.

$$\mathrm{GP}_{\mathrm{relay}}^{\mathrm{optimal}} = \max_{\substack{j_i \in \{1,\ldots,7\}, \\ i=1,\ldots,4}} \sum_{i=1,\ldots,4} \mathrm{GP}_{\mathrm{relay}}(i, j_i) \tag{5.14}$$

5.2.3 Numerical Results

Following the equations obtained in Section 5.2.2 the total throughput capacity of the direct and relay communication for the wireless network model under consideration is examined. Setting the PER requirement to 10%, the terminal separations d and d' for a given pathloss coefficient α and Noise N are varied. The transmission power P_t adjusts itself accordingly until it reaches the maximum limit. To avoid unbalanced scenarios such as a near terminal to the AP transmitting at full rate while the outer terminals achieve no throughput at all, a fair resource sharing mechanism is introduced. Hence, all terminals get a certain minimum amount of throughput; otherwise the total throughput is declared to be zero.

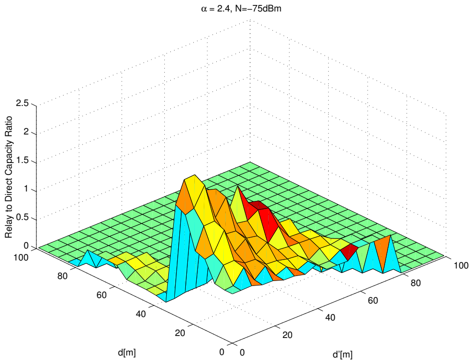


Figure 5.3: Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -75 dBm

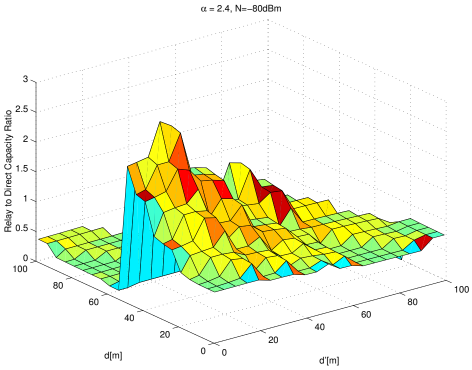


Figure 5.4: Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -80 dBm

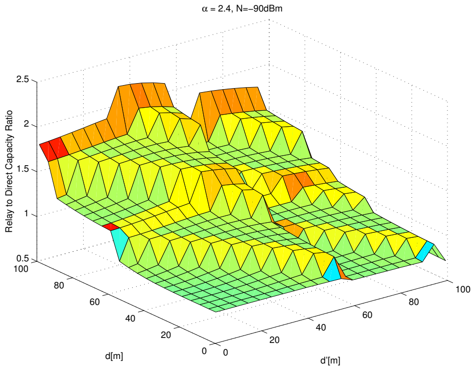


Figure 5.5: Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -90 dBm

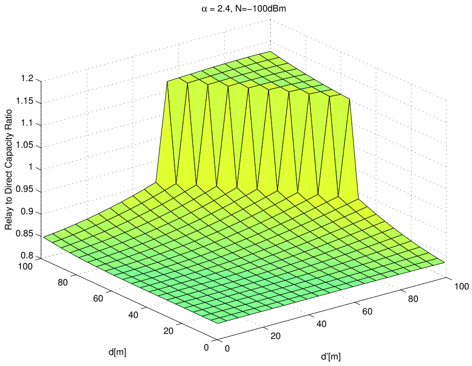


Figure 5.6: Capacity of a transmission power limited system for $\alpha = 2.4$ at N = -100 dBm

The following graphs (Figure 5.3 to Figure 5.18) give an overview of the optimal total throughput ratio of relaying to direct communication in a cell for various values of α and noise; i.e., $\mathrm{GP}_{\mathrm{relay}}^{\mathrm{optimal}}/\mathrm{GP}_{\mathrm{direct}}^{\mathrm{optimal}}$. Ratios with values larger than 1 indicate that relaying performs better than direct communication.

The result of the analysis indicates that the benefit of relaying becomes vivid when terminals are at the periphery of the cell and they have almost lost connection with the access point. This is due

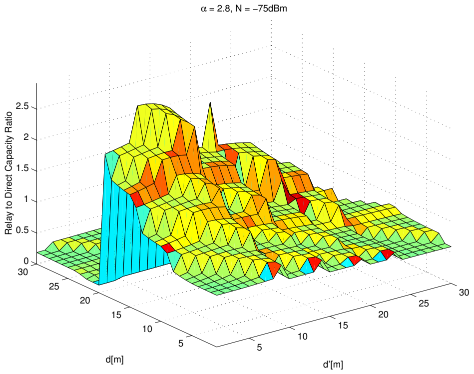


Figure 5.7: Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -75 dBm

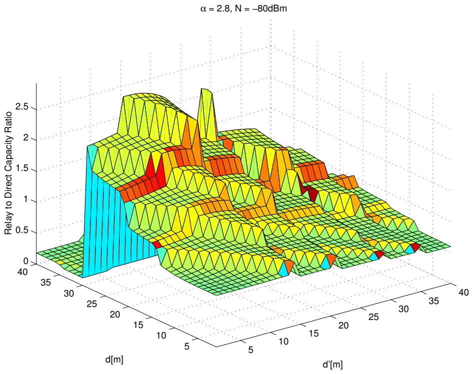


Figure 5.8: Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -80 dBm

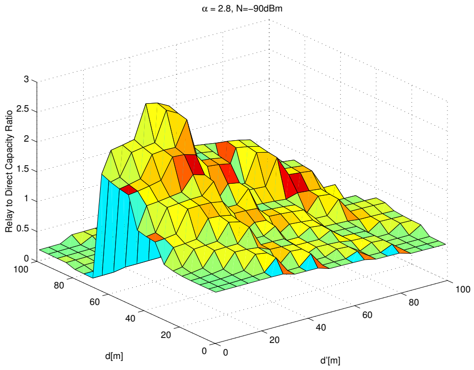


Figure 5.9: Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -90 dBm

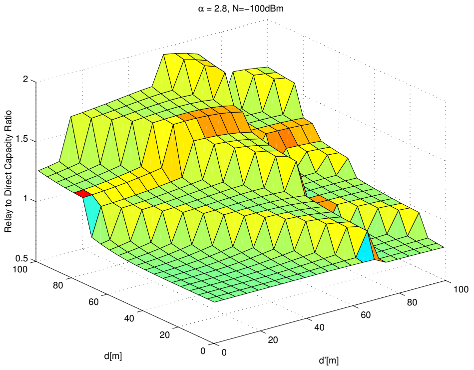


Figure 5.10: Capacity of a transmission power limited system for $\alpha = 2.8$ at N = -100 dBm

to the fact that the terminals have limited transmission power and they can only communicate upto the maximum allowable transmission power and the lowest possible modulation rate. Otherwise their packet error rate will be too high to sustain the communication. The performance improvement varies for different pathloss coefficients and noise figures. For smaller α values, the outer terminals are equally capable of sending with the highest transmission rate without the need for an intermediate terminal. But for higher α values, terminals can directly communicate with the highest possible data rate only when they are very close to the AP. Thus, relaying becomes prominent when the pathloss coefficient is large and when the terminals are very far from the AP.

In transmission power limited systems, noise can also influence the performance of the network. As shown in the figures, relaying outperforms direct communication when the network becomes noisy. For example, for N= -100 dBm the range of communication is relatively large. As the noise figure increases, the communication range decreases. However, the advantage of relaying is remarkable even for the reduced range. For different combinations of α and noise, it is possible to achieve more than 250% capacity increase in the network.

It is observed from the figures that there is a sharp decrease in the relaying throughput as the range of communication increases. This is due to the conservative scheduling requirement, i.e., the relaying terminal should transmit its own and the relayed traffics with the fastest transmission rate. However, this is not always possible when the outer terminals are close to the relaying terminal but the relaying terminal is relatively far from the AP. In such cases, the outer terminals can perform better when they communicate directly to the AP than using the relaying terminal. There are also cases where a single terminal gets the maximum share of the total throughput and the rest performs with very small bit rate. Although such cases can increase the total throughput and the capacity ratio considerably, they are avoided not to have unnecessarily unbalanced throughput distribution in the system. The sharp edges shown in the figures are the results of such cases as well. Moreover, there are certain configurations in which direct communication does not achieve any throughput at all, while relaying still performs well. In a sense, this means that relaying is able to extend the coverage of a cell – a not surprising result, which is however not the primary goal of the analysis.

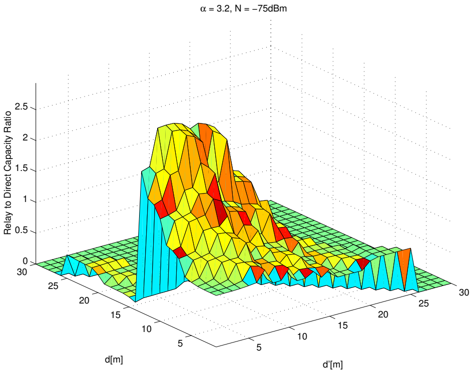


Figure 5.11: Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -75 dBm

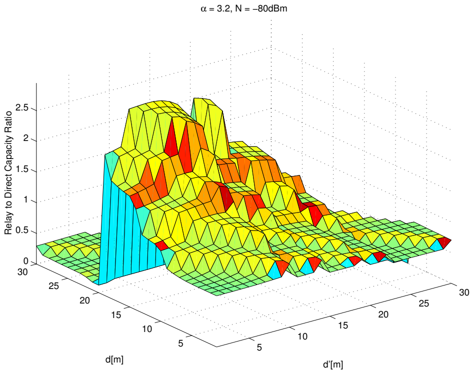


Figure 5.12: Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -80 dBm

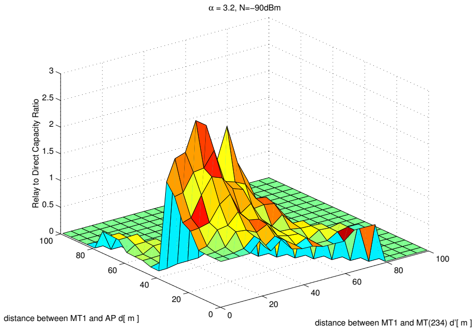


Figure 5.13: Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -90 dBm

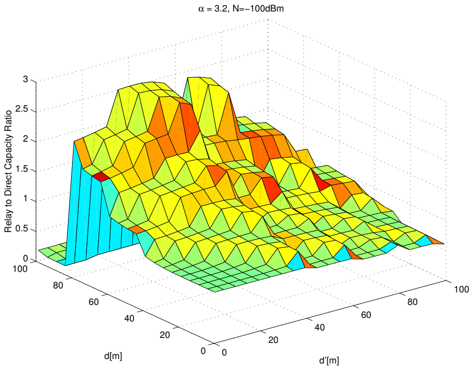


Figure 5.14: Capacity of a transmission power limited system for $\alpha = 3.2$ at N = -100 dBm

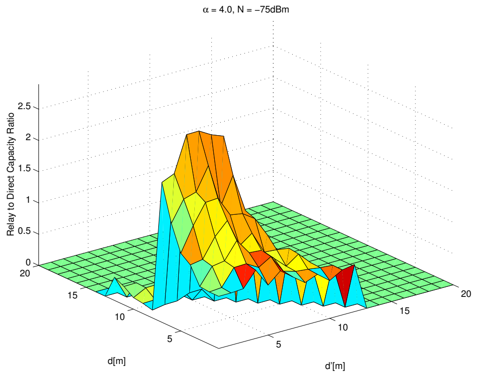


Figure 5.15: Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -75 dBm

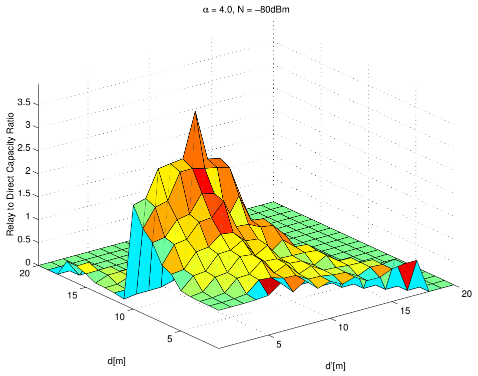


Figure 5.16: Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -80 dBm

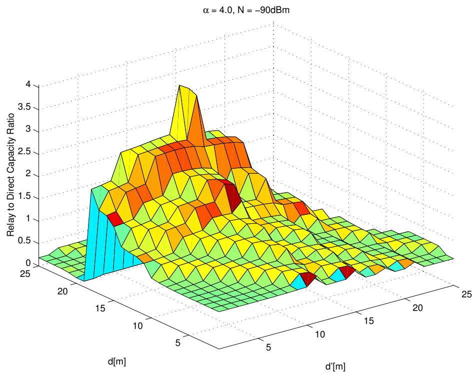


Figure 5.17: Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -90 dBm

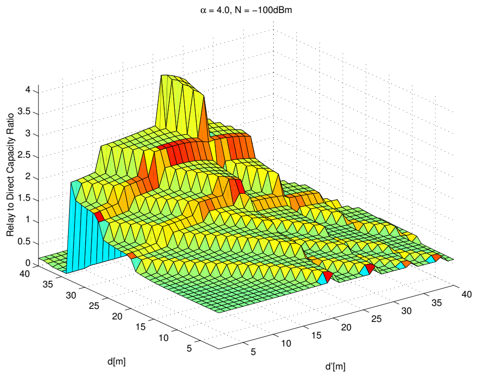


Figure 5.18: Capacity of a transmission power limited system for $\alpha = 4.0$ at N = -100 dBm

5.2.4 Summary

The performance of direct and relaying communication wireless cellular network is analyzed in this section. Noise and limited transmission power are taken to be the main limiting factors for capacity in the system. It is shown that relaying does indeed have the potential to improve upon the direct communication system when terminals are far away from their access points. The degree of improvement depends on the location (relaying performs better over large distances), the noise (relaying performs better in environments with a lot of noise), and the path-loss coefficient (relaying performs better when attenuation is large). This combination should make relaying particularly attractive in semi-static, noisy indoor environments, where a lot of traffic should be handled with few access points; examples for such environments are exhibition halls or temporary, ad-hoc installations with only few connections to a wired backbone.

Even though the model considered here is fairly simple, it allows some generalizations. For example, only a few numbers of terminals were considered, nevertheless, the results directly scale to large numbers as the relative share of time used by the terminals is the same in both direct and relaying communication approaches.

5.3 Interference Limited System

One of the main cause of transmission errors in interference limited systems is interference. In a direct communication scenario, every terminal is communicating directly with the access point. While this does not cause problems for terminals that are close to the base station, terminals that are far away (called “far” terminals) need to use an overproportionally high transmission power (due to the non-linear decay of received power over distance). This high transmission power, in turn, will create interference in the co-channel cells. Even though such cells will be some distance away (using typical frequency reuse schemes), the fact remains that reducing the transmission power of terminals close to the edge of cell holds a big promise for reducing interference and, hence, improving capacity.

However, directly reducing transmission power will result in a much higher error rate when attempting to communicate directly with the access point. On the other hand, communication via a relaying terminal is still possible. But this relaying terminal now has to transmit a higher traffic load (both the relayed and its own traffic), which is in general not possible at the bit rate of the modulation it is currently using. To overcome this, a “faster” modulation with a high bit rate needs to be used. Normally a higher bit rate incurs a higher PER as compared to the lower one at the same level of noise and interference. But since the far terminals reduce their power due to relaying, the level of interference is reduced and the use of high bit rate can actually be feasible.

In the previous section, it is shown how the capacity of cellular wireless system is influenced by limited transmission power, assuming negligible interference from neighboring cells. In this section, the tradeoff between lower interference by reduced transmission power at the border of cells on the one hand and increased error rate due to the need to increase data rate in the interior of a cell on the other hand is analyzed.

5.3.1 Model

For analytical investigation of relaying a two cell model, each cell with an AP and two mobile terminals, is considered. To simplify the analysis, a symmetric layout of access points and cells has been chosen, which is shown in Figure 5.19: all six terminals are located on a single line, the distance between terminals within the cells is indicated by d, the distance between the two outmost terminals of each cell is D, which is actually the separation of the cells. As the interest of this section lies on the effect of interference, both cells are assumed to use the same communication channel.

Again here, the terminals are assumed to have unlimited data to transmit with the maximum possible data rate to the AP. Let the inner terminals M_1 and M_4 use transmission power P_T and the outer terminals M_2 and M_3 use transmission power P'_T. A simple pathloss model is used to relate the

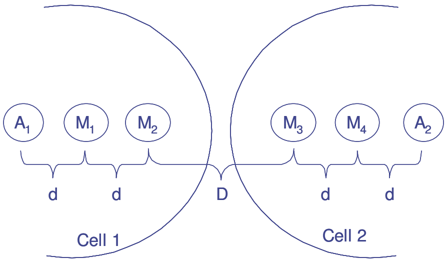


Figure 5.19: Scenario description

	Cell 1	Cell 2
Time slot 1	$M_1 \rightarrow A_1$	$M_3 \rightarrow A_2$
Time slot 2	$M_2 \rightarrow A_1$	$M_4 \rightarrow A_2$

Table 5.4: Schedule for direct communication

tranmitted and received power P_T and P_R, and it depends only on the terminal separation distance x and pathloss coefficient α, i.e, $P_R = \frac{P_T}{x^\alpha}$ for $x > 0$.

The actual packet error rate results from the received power, the noise N and the interference from other transmissions in the same channel. In this model, adjacent channel interference is ignored. Thus, the signal to noise and interference ratio is give by:

$$SINR = \frac{P_R}{N + I} \tag{5.15}$$

The exponential curve fitting discussed in Section 5.2.1 is also applied here to determine the corresponding PER for a given SINR and modulation ranging from 6 to 54 MBit/s. Similar to the limited transmission power case, all terminals have unlimited traffic to send and they are assumed to transmit at the maximum possible data rate to the access point.

5.3.2 Analysis

Direct Transmission

Initially a communication schedule needs to be determined in order to facilitate the computation of the interference in the two cells and the throughput at the AP. The uplink phase considered here is scheduled using two time slots, one for the inner and another for the outer terminals. To keep the symmetry between the two cells, inner and outer terminals of the two cells operate in alternate slots. These terminals, independent of each other can use any possible modulation to send their traffic. The resulting schedule is shown in Table 5.4.

Based on the schedules, the actual SINR values can directly be computed for transmitting from M_1 to A_1 and from M_2 to A_1 as:

	Cell 1	Cell 2
Time slot 1	$M_1 \rightarrow A_1$ (fast modulation)	$M_3 \rightarrow M_4$
Time slot 2	$M_2 \rightarrow M_1$	$M_4 \rightarrow A_2$ (fast modulation)

Table 5.5: Schedule for relay communication

$$SINR_{M_1 \rightarrow A_1} = \frac{\frac{P_T}{d^\alpha}}{N + \frac{P_T}{(2d+D)^\alpha}} \tag{5.16}$$

$$SINR_{M_2 \rightarrow A_1} = \frac{\frac{P_T}{(2d)^\alpha}}{N + \frac{P_T}{(3d+D)^\alpha}} \tag{5.17}$$

Note that the derivation of one cell is symmetrical to the other one and hence, only one cell is considered. Once the SINR values of each transmission is computed, the PERs can easily be determined from Equation (5.3) for the employed modulation type i. The throughput for a particular terminal can then be approximated as $GP = (1 - PER)NBR$.

However, since the two terminals in the cell have to share the channel with each other, the nominal date rate needs to be divided by two (without considering a downlink phase, which would further scale down the throughput). Thus, the throughput is

$$\mathrm{GP}_{M_1 \rightarrow A_1} = (1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{M_1 \rightarrow A_1}), j))\frac{\mathrm{NBR}_j}{2} \tag{5.18}$$

$$\mathrm{GP}_{M_1 \rightarrow A_1} = (1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{M_1 \rightarrow A_1}), j))\frac{\mathrm{NBR}_j}{2} \tag{5.19}$$

Cell 1's total throughput is the sum of throughput of terminal 1 and 2.

Relay Communication

The schedule for this setup is slightly different: On the one hand, only two time slots can be used to communicate the traffic of both mobile terminals within a cell to the access point. On the other hand, the relaying stations M_1 and M_4 should transmit both their own traffic as well as the relayed traffic of M_2 and M_3, respectively. Evidently it is not possible to use the same amount of time and the same modulation as in the previous case (when both stations transmit at a bandwidth close to the maximum possible bandwidth of a chosen modulation). Hence, to transmit from the relay station to the access point, a “faster” modulation has to be used. Here, we assume that for this communication link, a modulation at least twice as fast as the communication between normal mobile terminal and relay terminal is used. Thus, the actual schedule would look like as in Table 5.5.

Since it is already assumed that all terminals have unlimited data to transmit, the relay terminals always use the faster, more error-prone modulation. But in reality, outer terminals may not always have traffic to send and the relay terminal can be flexible in choosing a modulation rate. Adding a mechanism that uses faster modulation between relay and access point only when necessary should again improve the error behavior of this link. Such mechanism, however, could imply additional signaling overhead.

Based on the schedule, the actual SINR values for transmitting from M_1 to A_1 and from M_2 to M_1 can be computed respectively as

$$SINR_{M_1 \to A_1} = \frac{\frac{P_T}{d^\alpha}}{N + \frac{P_T}{(2d+D)^\alpha}} \tag{5.20}$$

$$SINR_{M_2 \to M_1} = \frac{\frac{P_T}{d^\alpha}}{N + \frac{P_T}{(2d+D)^\alpha}} \tag{5.21}$$

The main difference between the direct and the relay case is now apparent: lower transmission power is used close to the border of a cell, improving the interference situation. Yet on the downside, faster and more error-prone modulations need to be used close to the interior of a cell. Determining PER and throughput is also different from the direct case as two different modulations going over the different hops are used.

Consider first the throughput for terminal M_1, which uses modulation i with nominal bit rate NBR_i to transmit to the AP. Since M_1 is transmitting only for half the time, NBR_i has to be divided by two. Moreover, it is only part of the data that belongs to M_1 in the given time slot. Thus, the actual throughput of M_1 for transmitting its own data is

$$\mathrm{GP}^{relay}_{M_1 \to A_1}(i,j) = (1 - \mathrm{PER}(10 \log_{10}(SINR_{M_1 \to A_1}), i)) \frac{(NBR_i - NBR_j)}{2} \tag{5.22}$$

where j is the modulation index used by M_2 to relay its data to M_1. The throughput of M_2 depends on both the NBR_j it is using when communicating with M_1 and the NBR_i that M_1 is using to communicate with the access point. Similar to the direct case, M_2 can transmit only half the time and hence its NBR is reduced by half. The bandwidth from M_1 to A_1 should not limit the throughput M_2 since it is already assumed that $NBR_i > NBR_j$. More complicated is the error handling: a packet from M_2 needs to be transmitted successfully between M_2 and M_1 as well as between M_1 and A_1. Hence, the throughput from M_2 is:

$$\mathrm{GP}^{relay}_{M_2 \to A_1} = (1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{M_1 \to A_1}), i))(1 - \mathrm{PER}(10 \log_{10}(\mathrm{SINR}_{M_2 \to M_1}), j)) \frac{\mathrm{NBR}_j}{2}. \tag{5.23}$$

Just like in the direct communication case, these numbers would have to be downscaled correspondingly for inclusion of a downlink phase. However, as this needs to be done in both cases, the scaling factor cancels out when looking at the ratio of the total throughput in both cases and is hence not considered here. The total throughput at the AP is then the sum of the individual terminal throughput in the corresponding cell. The optimization problem, hence, lies in selecting the right modulation and transmission power such that the total throughput is maximized for any given scenario,

5.3.3 Numerical Results

The capacity of both direct and relay communication of the wireless network model is examined using numerical solutions of the equations derived in Section 5.3.2. The result of this numerical solution depends on the mobile terminal separation d and the intercell separation D for various values of transmission power P_T and path loss coefficient α as well as on the noise level N; the results are the SINR and the throughput for both near and far terminal using either direct or relay communication.

Figures from 5.20 to 5.23 give an overview of the ratio of the total throughput in a cell between direct and relaying communication. Ratios with values larger than 1 indicate a better relaying performance. For both communication paradigms, all terminals use the best combination of modulation types and transmission powers in order to achieve a maximum total throughput. As can be seen from the figures, the capacity ratio improves with increasing α. For example, an improvement of about 60% is achieved when the terminals are far away from the AP. Evidently, there are placements of terminals where direct transmission is superior and others where relaying communication is better.

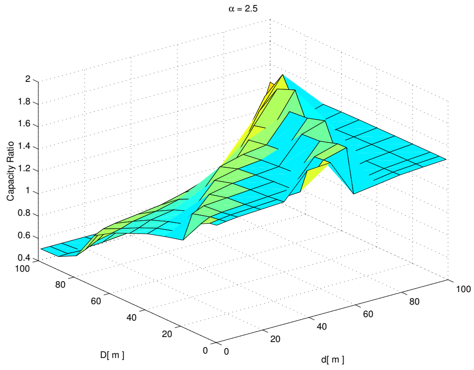


Figure 5.20: Ratio of total throughput for $\alpha = 2.5$

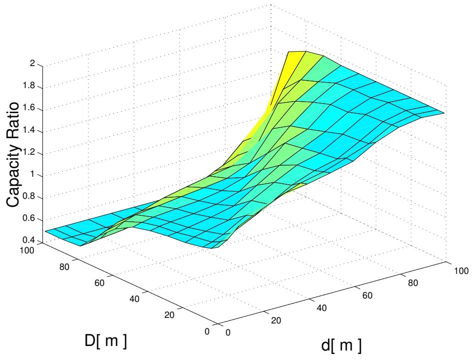


Figure 5.21: Ratio of total throughput for $\alpha = 3.0$

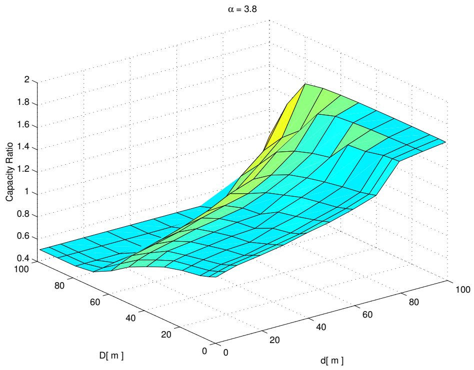


Figure 5.22: Ratio of total throughput for $\alpha = 3.8$

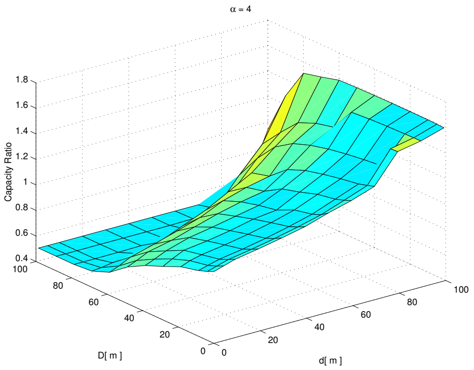


Figure 5.23: Ratio of total throughput for $\alpha = 4$

It is possible to conclude that relaying is particularly beneficial when considering scenarios where mobile terminals are far away from their AP, and for the case of co-channel cells located close to each other. While the last case is usually avoided with spatial reuse, it might occur more often when relaying within a cell which uses multichannel communication technology.

5.3.4 Summary

In this section, it is attempted to probe the potential of relaying in a simple scenario. The analysis and numerical results have shown that relaying can achieve considerably larger total throughput than direct communication in a TDMA-based wireless network. Critical parameters are the path loss coefficient α and the physical location of terminals. In this particular setup, best results were achieved when terminals were far away from their access points, yet the far terminals were still comparably close to each other. However, this means that the access points themselves were still separated by a considerable distance, as would be required be a spatial reuse scheme.

It is possible to look at the results under a different perspective: The particular strength of relaying is a setup where cells are comparably large, yet close together. Hence, if a certain throughput is required per cell, it is possible to place such cells closer together when relaying is used, as the spatial "protection" between cells need not be as large as with direct communication. This should prove particularly beneficial in scenarios where a lot of cells should be placed within a limited amount of space, e.g., office buildings, exhibition halls, etc.

5.4 Probability Estimation of Finding Relaying Terminals

In the previous sections, the potential advantage of relaying is analyzed for transmission power and interference limited systems. In both cases, it is assumed that there exists a relay terminal which is ready to transmit a relayed traffic. But it is important to know how often such relaying terminals are readily available in the network. This section presents the probability of finding relaying terminals in a cell by taking an analytic network model. The model analysis is based on a geometric construction, the result of which can be generalized for any link technology [1][32].

5.4.1 Model and Assumptions

Consider a simple cellular network with N terminals distributed within the cell and send their traffic to the AP which is located at the center. Far terminals can send traffic either directly or via a relaying terminal. An instance of such a model is shown in Figure 5.24.

Each terminal in the network supports a small number, k, of transmission modes $M_1, \ldots, M_k$. With each mode, a maximum transmit radius (for some maximum PER), $d_1, \ldots, d_k$, is associated, where $d_i < d_j$ for $i < j$; d_k is assumed to define the cell radius. Each transmission mode M_i also

---

[1] This section is a joint work of the Swidisch Institute of Computer Science and Technical University Berlin

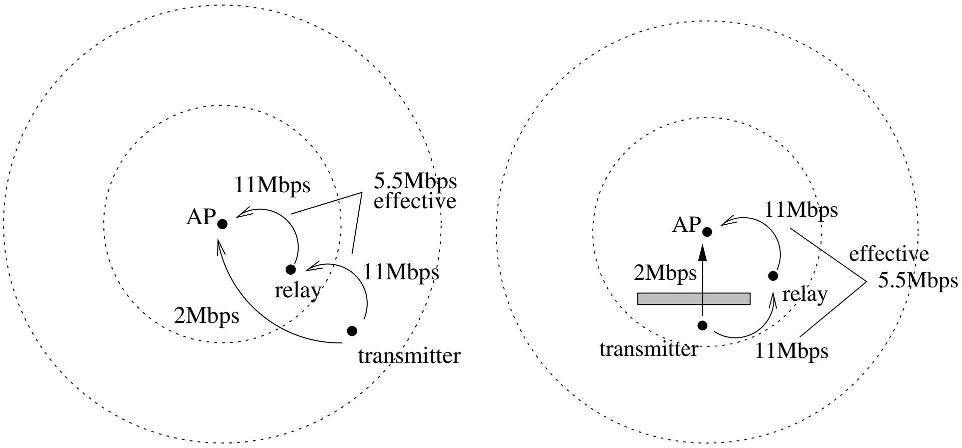


Figure 5.24: A simple example of intercell relaying. Relaying can be beneficial for communicating over a longer distance or around an obstacle

has associated with it some transmission "cost" C_i. It is assumed that there is a direct relationship between distance and cost, thus $C_i < C_j$ for $i < j$ (the higher the cost, the longer distances can be covered). Relaying here is defined as partitioning a transmission into a sequence of transmissions with a lower overall cost than the corresponding direct transmission. It is required that the partitioned transmissions are independent to each other.

To obtain the probability that for a given terminal, there is another terminal in such a position that relaying reduces the total cost of communication, a relay region is defined as shown in Figure 5.25. The relay region $R_t(x, y)$ is the relay region of transmitter t at position (x, y), as the region in which a station must be located in order to act as a relay for t.

The relay region is non-empty only for $1 < r \leq \min(D, 2)$. For $0 < r \leq 1$, the transmitter itself is within relay radius of the access point. Although there may be stations that can act as a relay for the transmitter, there is no cost benefit obtained by relaying since the station can already communicate with the access point at the lower cost. The relay region is therefore defined to be empty. For $2 < r$, the transmitter is too far away from the access point to communicate with it using a single relay, so the relay region is again empty. For $D < r$, the station is outside the cell. For simplicity, only two transmission rates are shown in the example figures.

Without loss of generality, let the access point be located at $(0, 0)$ and the transmitter t at $(r, 0)$; r then denotes the distance between them. In order for a station to act as a relay for a transmitter, the distance between the relay and the transmitter must be less than the relay radius; otherwise, the transmitter cannot communicate with the relay using the lower cost (smaller transmission radius) transmission mode. Similarly, the distance between the relay and the access point must also be less than the relay radius; otherwise the relay cannot communicate with the access point at the lower cost transmission mode. A similar geometric approach is also used in [83] to find the probability that any two nodes can be connected by a two-hop path in an N-node randomly distributed mobile networks.

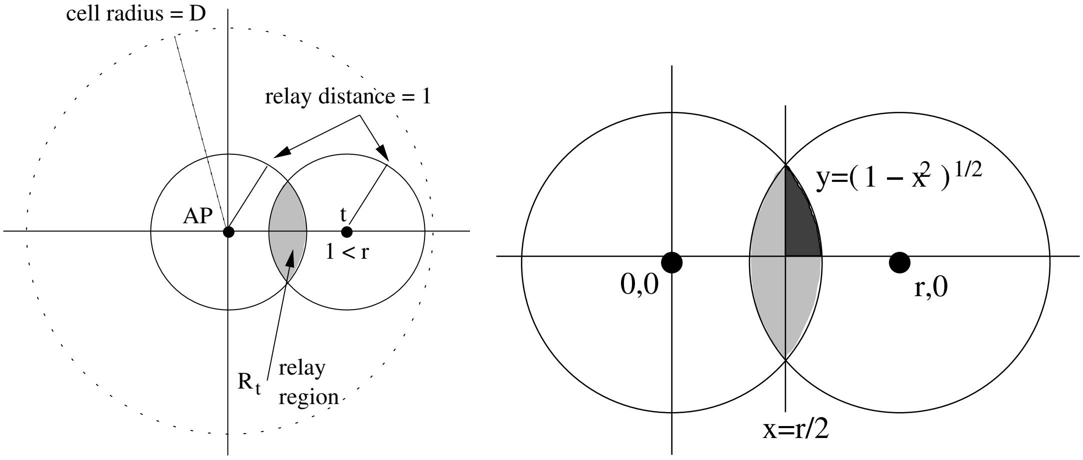


Figure 5.25: Relay region (left); Detail (right)

5.4.2 Numerical Results

From the model and assumptions described above, the probability that, given a transmitter t, at least one of the other $N-1$ stations is in its relay region $R_t(r)$ is given by:

$$P_{n\geq 1}(r) = 1-(1-p)^{N-1} = 1-\left(1-\frac{A(r)}{\pi D^2}\right)^{N-1} \tag{5.24}$$

where $A(r)$ is area of the relay region. And the expected proportion of stations in the network which can take advantage of relaying is given by:

$$\begin{aligned} P(N,D) &= \lim_{\Delta r\to 0} \sum_{i=0}^{\lfloor \frac{\min(D,2)}{\Delta r} \rfloor} \left(P(\text{terminal in ring from } i\Delta r, \Delta r \text{ wide})\ P_{n\geq 1}(i\Delta r)\right) \\ &= \int_{r=1}^{\min(D,2)} \frac{2r}{D^2} P_{n\geq 1}(r)\mathrm{d}r = \int_{r=1}^{\min(D,2)} \frac{2r}{D^2} \left(1-\left(1-\frac{1}{\pi D^2}A(r)\right)^{N-1}\right) \mathrm{d}r \end{aligned}$$

Note that the detailed analysis can be found in [32].

The behavior of $P(N,D)$ is illustrated for various values of N and D in Figure 5.26. Strikingly, the proportion of stations that can take advantage of a relay grows rapidly as the node density increases, starting from less than 20% for four node networks to reach an upper bound of 75 %. This shows that the probability of relaying grows when it is most needed, in densely populated networks.

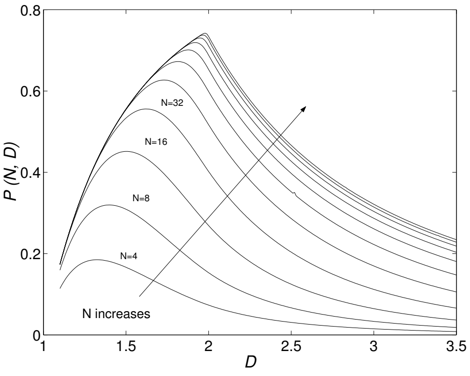


Figure 5.26: Behavior of $P(N, D)$ for various D as N becomes large (graph shows curves for $N = 4, 8, 16, \ldots, 4196$)

Chapter 6

Approach

6.1 Introduction

The analysis of both transmission power and interference limited systems reveal that terminals close to the access point can reasonably communicate without any need for relaying. Relaying becomes beneficial to the system when terminals are far from the access point and when co-channel cells are close to each other, i.e., the frequency reuse distance is small. In the analysis, it is assumed that there is always an intermediate terminal available which is ready to relay for other terminals. But this may not always be the case if the network is sparse. Users may not also be willing to assist in relaying to save their battery or from fear of security problems like viruses carried by that traffic [79]. Thus, availability of terminals is also an important issue.

The probability analysis shown in Section 5.4 is done under the assumption that terminals are always willing to relay, the willingness issue is left for simplicity. The probability results indicate that as the network density grows, the proportion of terminals that take advantage of relaying also grows. The network models used in the analysis are rather simple, nevertheless, they reasonably confirm the potential of relaying to improve capacity of system.

Putting the separately studied systems together, the performance of the wireless system is influenced by both the strength of the transmission power, noise and interference. The motivation now is to make use of the indicated potential to the real system. There should be a mechanism to decide if relaying is necessary, in the first place. The other issues are how to select the right relaying terminal, how to choose an optimal transmission power and data rate and how to share the available resource fairly to all transmitting terminals. This chapter is dedicated to describe the unique approaches used in this work to improve the system capacity.

The remaining of this chapter is organized as follows. Section 6.2 discusses how transmission power and rate can optimally be selected. Section 6.3 presents methods of selecting relaying terminals in the network. Fairness constraints in the network and ways of scheduling the communication are described in Section 6.4 and 6.5. Finally Section 6.7 gives an example scenario that summerize the proposed mechanisms.

6.2 Transmission Power and Rate Adaption

The potential benefit of relaying depends on the data rates that can be realized between relayed terminal, relaying terminal, and access point. On the one hand, faster data rates can be selected when the distance between two communicating terminals is reasonably small. But since faster data rates are prone to high packet error rate [61], it may be necessary to use higher transmission power. Use of high transmission power however has a potential danger to cause interference. On the other hand, slower data rates can be selected to reduce the packet error. But when the communication distance increases, lowering the data rate may not be enough, transmission power needs to be increased to remain in the allowable packet error rate range. In either ways, there should be an optimal choice of data rate and transmission power to best achieve a larger throughput.

The effective data rate between any two terminals can be determined based on their channel gain and a target packet error rate. Here, the channel gain is assumed to be derived from the distance between terminals. For each data and a target packet error rate, the required transmission power is computed as follows:

$$P_{tx} = 10^{\frac{1}{10}(\mathrm{SINR}+\mathrm{N}+\mathrm{PL})} \tag{6.1}$$

where P_{tx} is the transmission power in mW, SINR is obtained from the approximately known relationship between signal-to-noise and packet error rate [61], N is the ambient noise figure and PL is the pathloss as described in Section 3.2:

$$L = 46 + 10 \cdot \alpha \cdot \log_{10}\left(\frac{\mathrm{d}}{\mathrm{m}}\right) \tag{6.2}$$

Any data rate that requires more than a maximum allowable power (here, 200 mW) or that does not match minimal required receiver sensitivity is ruled out. As a result, the optimal data rate for this pair of terminals is obtained. The smallest transmission power is used that still meets the target PER for this data rate. It is assumed that the fading in the network is for over a longer period and it has a uniform effect.

An example of such an effective-data-rate selection is shown in Figure 6.1 for 1% target PER. The shape of this figure is the primary justification for hoping that relaying will actually improve capacity: For many distances, the achievable data rate over half the distance is more than twice the data rate over the full distance. This observation holds for other α and target PERs as well.

It should be pointed out that this transmission power/rate adaptation introduces another tradeoff: By switching to higher rates over shorter distances, it can be necessary to increase the transmission power. This increased power in turn increases interference in neighboring cells. In addition, since

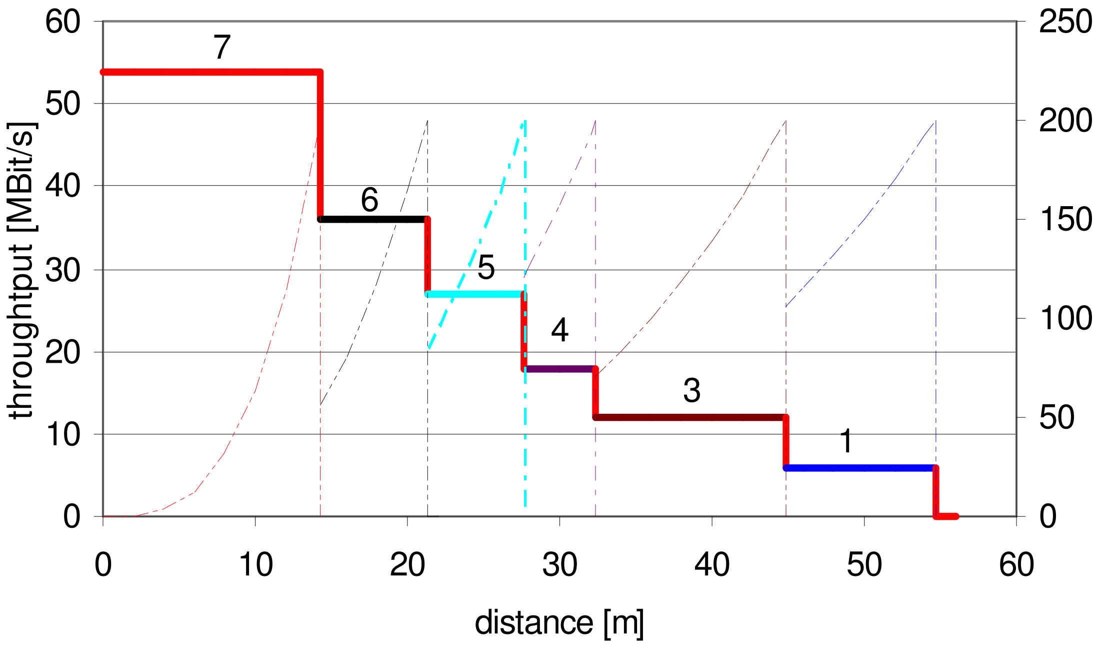


Figure 6.1: Effective data rate and transmission power with respect to distance for 1 % PER and $\alpha = 3.2$

terminals close to the AP may transmit their own as well has relay traffic, the energy consumption of relay terminals may be high. This might sound like relaying terminals being disadvantaged compared to relayed terminals in energy/battery consumption. Such issues are discussed in detail in Chapter 10.

6.3 Routing

It is evident that relaying is used when there are other terminals which can serve as an intermediate terminal to relay traffic coming from terminals which are relatively far from the access point. Thus, relaying is envisaged to be more practical in an urban area where there is dense population. This section addresses how the intermediate terminals can be selected and traffic routing is done if agreed upon the availability such terminals.

As mentioned in Section 3.3, the hexagonal cellular structure is considered in an area where there is a wireless communication service and relaying is applicable. Each cell consists of an access point at the center and a number of terminals which are randomly distributed within the cell. The radius of a cell is determined by the maximum possible distance a terminal can communicate with the AP using the slowest data rate for a given target PER and pathloss coefficient. Thus, there is no coverage problem within and at the edge of each cell. The question now is to determine the number of hops needed to relay the far terminal's traffic to gain a reasonable capacity improvement with relatively less relay management overhead.

Consider an example of a relaying setup shown in Figure 6.2. Let the end-to-end throughput (from MT1 to AP) be designated as $\Upsilon(d,k)$, where d represents the distance between the source and destination terminal and k is the number of hops needed by $k-1$ intermediate (relay) terminals to finally reach the AP. For the direct communication case, the outer terminal MT1 uses only one hop for its communication and it uses the whole communication time allocated as shown in (b). The corresponding throughput becomes $\Upsilon(d,1)$.

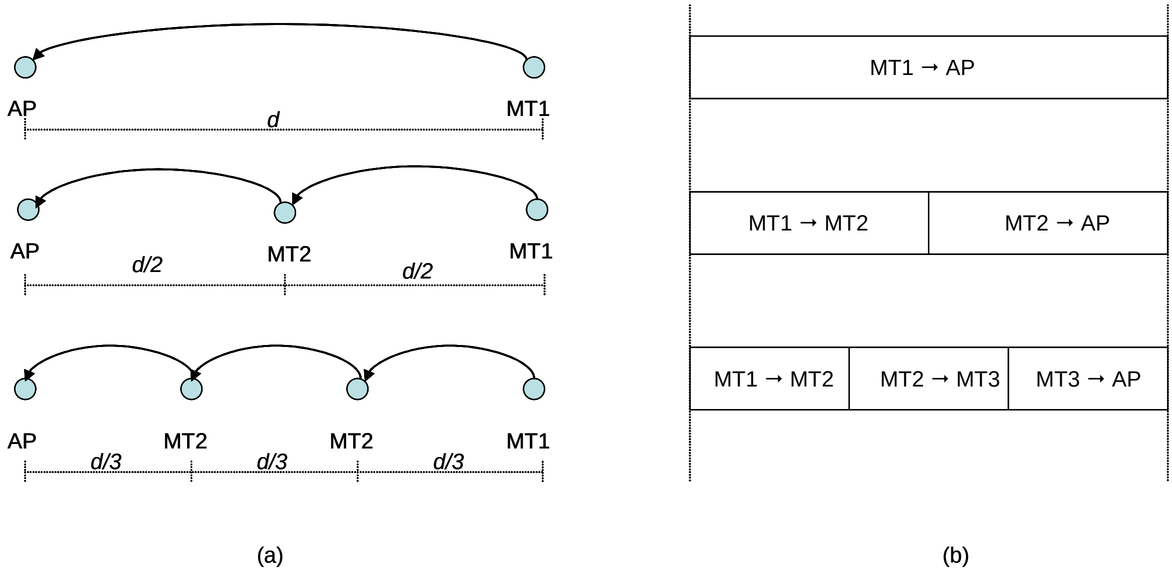


Figure 6.2: (a) Number of hops needed to send traffic from the far terminal to the AP; (b) Communication time shared among the terminals involved when relaying in one-frequency is used

For the relaying case, let us assume that the relay terminals are placed in an optimal positions, i.e., each terminal covers the same distance to send traffic to the next terminal. If interference is not taken into account, each terminal will have the same data rate and share the communication time equally as indicated in Figure 6.2(b). The end-to-end throughput for two-hop or single-relay hop becomes $\frac{1}{2}\Upsilon(\frac{d}{2},1)$. Similarly for the three-hop, the throughput is $\frac{1}{3}\Upsilon(\frac{d}{3},1)$. In general, for k number of hops the end-to-end throughput is

$$\Upsilon(\mathrm{d},\mathrm{k}) = \frac{1}{k}\Upsilon\left(\frac{\mathrm{d}}{\mathrm{k}},1\right) \quad \text{for } k>1 \tag{6.3}$$

Figure 6.3 shows an example of end-to-end throughput for a terminal in a cell sending its traffic to the AP through different number of hops. The benefit of relaying to increase the throughput becomes clear as the terminal goes far from the AP. It is also possible to see from the figure how relaying is also beneficial for coverage extension and for cells with relatively larger radius. The end-to-end throughput gain when shifting from direct communication to two-hop or single-relay hop is about 123%, for a

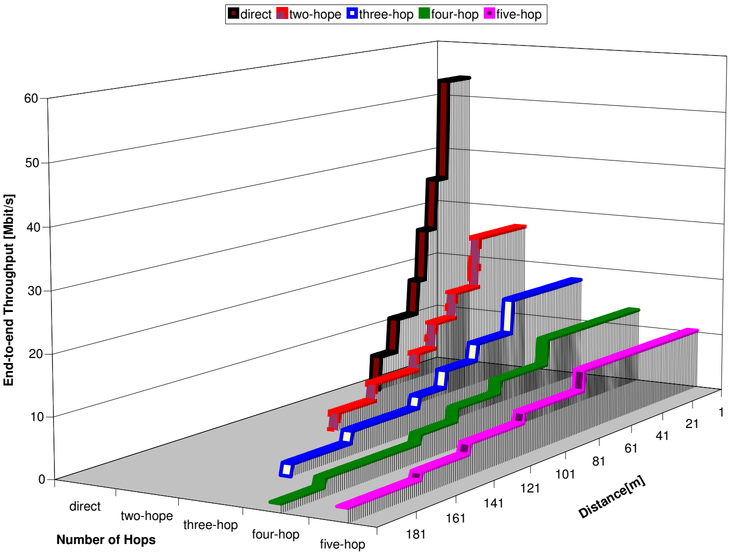


Figure 6.3: End-to-end throughput as a function of source-destination distance for k number of hops; 1% PER and $\alpha = 3.2$

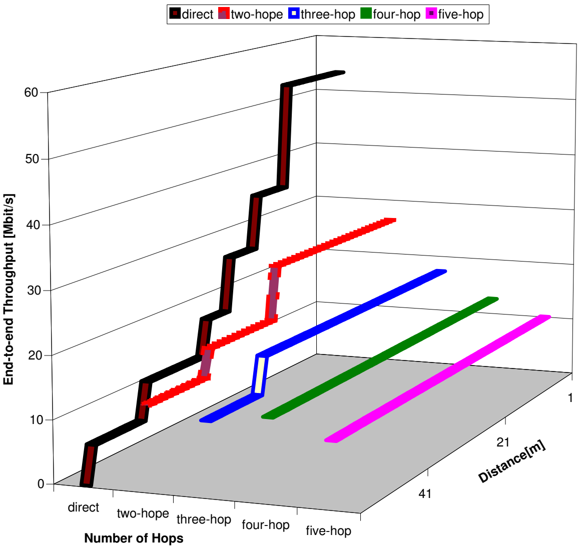


Figure 6.4: End-to-end throughput for a cell radius of d; 1% PER and $\alpha = 3.2$

far terminal placed at 50m. For the same position, the gain when shifting from two-hop to three-hop is rather -12% and from three- to four-hop is again -11.7%. But the negative does not necessarily mean that multi-hop communication is not beneficial. For another position which is closer to the AP, say 25m, switching from direct to two-hop does not change the end-to-end throughput and if more than two-hop is used, a decreasing effect is observed. Therefore, for cellular structure whose radius is limited by the distance of the farthest terminal from the AP, which is sending with maximum allowable transmission power and slowest data rate, it is reasonable to assume that two-hop or single-relay-hop suffice to show the potential of relaying for improving the system capacity. The other rationale behind the choice of single-relay-hop is to reduce the cumulative impact of link and physical layer overhead at each hop and to reduce the problem of additional complexity, frequency switching management and overhead that may be required to support multihop relays.

6.3.1 Relay Selection

Though relaying can potentially improve the individual link throughput, terminals may not always need to relay their traffic as sending directly to the destination can sometimes be beneficial as well. Selecting the right terminal whenever relaying is needed, thus, plays crucial role in studying the network capacity.

The relaying decision is done by first computing the maximum achievable throughput when a terminal is communicating directly with the access point in a given time slot. This throughput is then compared with the throughput that can possibly be achieved when the terminal is alternatively using each of available potential relaying terminals in the network for the same time slot as in the direct case. If any of the alternative routes provide a higher throughput than the direct one, the terminal chooses relaying. Otherwise, the terminal communicates directly with its access point.

After the relaying decision is made, the relay terminal selection is as follows: Suppose a far terminal MT has data size y and this data is directly transmitted to the AP in time slot t_{direct} length as shown in Figure 6.5. Suppose again that there exist a ready-to-relay candidate terminal r_{MT} for which to decide whether to relay via it or not. Assuming that all terminals fairly share the communication time, in order to use r_{MT}, we require for a candidate terminal relay:

1. data y can be sent from MT to r_{MT} in time slot t_1
2. relayed data y can be sent from r_{MT} to AP in time slot t_2
3. $t_1 + t_2 = t_{\mathrm{direct}}$
4. $\mathrm{GP}_{\mathrm{relay}} > \mathrm{GP}_{\mathrm{direct}}$

where $\mathrm{GP}_{\mathrm{direct}}$ and $\mathrm{GP}_{\mathrm{relay}}$ are the effective data rates for the direct and relay transmission, respectively. These data rates are computed as follows.

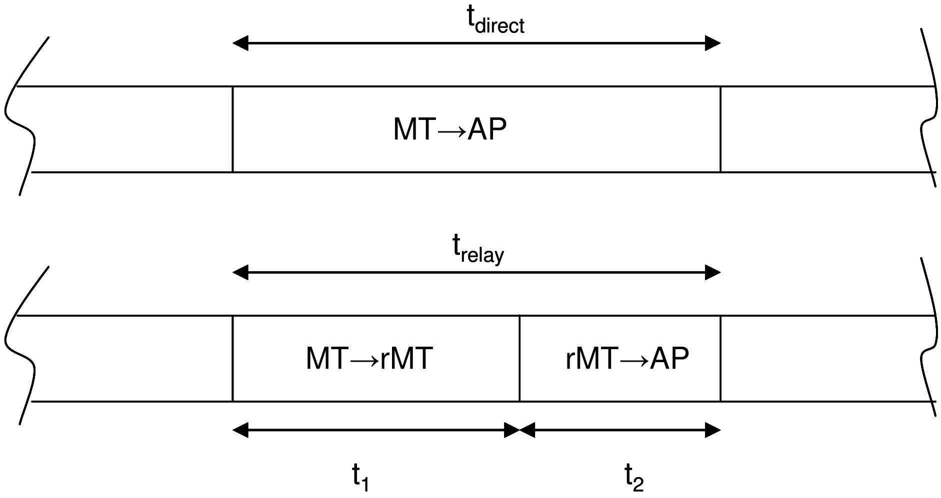


Figure 6.5: slot for direct and relay transmission

1. **Direct case**

   The effective data rate for the direct transmission is the correctly transmitted data for a given time, i.e

$$\mathrm{GP}_{\mathrm{direct}} = (1 - \mathrm{PER}_{\mathrm{MT}\rightarrow\mathrm{AP}})\mathrm{Rate}_{\mathrm{MT}\rightarrow\mathrm{AP}} \tag{6.4}$$

   where $\mathrm{PER}_{\mathrm{MT}\rightarrow\mathrm{AP}}$ is the packet error rate from MT to the AP.

2. **Relay case**

   When splitting up the direct time slot t_{direct} into two sub-slots, the question arises how to choose their lengths t_1 and t_2 such that the overall rate is maximized. Leaving aside the error rates on both links for the moment, a simple consideration shows that the data transmitted over both links should be equal to maximized throughput – otherwise, there would be too much or too little data arriving at the relay terminal. Hence:

$$\begin{aligned} t_1\mathrm{Rate}_{\mathrm{MT}\rightarrow r_{\mathrm{MT}}} &= t_2\mathrm{Rate}_{r_{\mathrm{MT}}\rightarrow\mathrm{AP}} \\ \Leftrightarrow \frac{t_2}{t_1} &= \frac{\mathrm{Rate}_{\mathrm{MT}\rightarrow r_{\mathrm{MT}}}}{\mathrm{Rate}_{r_{\mathrm{MT}}\rightarrow\mathrm{AP}}} \end{aligned} \tag{6.5}$$

   This means that the $t_1\mathrm{Rate}_{\mathrm{MT}\rightarrow r_{\mathrm{MT}}}$ amount of data can be transported from the far terminal to the AP in $t_{\mathrm{direct}} = t_1 + t_2$, resulting in a total throughput of

$$\frac{t_1 \text{Rate}_{\text{MT}\to r_{\text{MT}}}}{t_1 + t_2} = \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{1 + \frac{t_2}{t_1}} \overset{\text{Eq. (6.5)}}{=}$$

$$\frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{1 + \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{\text{Rate}_{r_{\text{MT}}\to \text{AP}}}} = \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}} \text{Rate}_{r_{\text{MT}}\to \text{AP}}}{\text{Rate}_{\text{MT}\to r_{\text{MT}}} + \text{Rate}_{r_{\text{MT}}\to \text{AP}}}$$

Taking the error rates on each link into account, the relay goodput is then:

$$\text{GP}_{\text{relay}} = (1 - \text{PER}_{\text{MT}\to r_{\text{MT}}}) \cdot (1 - \text{PER}_{r_{\text{MT}}\to \text{AP}}) \cdot \frac{(\text{Rate}_{\text{MT}\to r_{\text{MT}}})(\text{Rate}_{r_{\text{MT}}\to \text{AP}})}{(\text{Rate}_{\text{MT}\to r_{\text{MT}}}) + (\text{Rate}_{r_{\text{MT}}\to \text{AP}})}$$

Finally, among the candidate relay terminals, the one with maximum GP_{relay} is selected as a relaying terminal for the MT in consideration. Note that the relaying decision does not affect the entire schedule but only the individual time slots.

The relay selection mechanism is an *exhaustive search* which involves the entire terminal in the network. Alternatively, *sector based search* can also be used to reduce the computation.

Exhaustive Search Consider a cell with an access point and n terminals. For a terminal to decide whether relaying is beneficial, it searches the $n-1$ terminals by computing the achievable throughput when it is directly communicating and when it is relaying via each of the $n-1$ terminals. If the throughput obtained for relaying is larger than the direct, the terminal switches to relaying. Among the potential relaying terminals, the one which results in the maximum throughput is chosen as the relaying terminal.

Sector Based Search Here the computation is similar to the exhaustive case but the difference is only part of the cell is searched instead of the entire cell. As discussed in Section 3.3.3, mobile terminals are identified not only by the AP they belong to and their x and y positions but also by the sector they are associated within the hexagonal cell. An example of such sector based association is shown in Figure 6.6. Thus, terminal in sector s searches for a relaying terminal only in $\{s-1, s, s+1\}$. In other words the search is in half of the cell. This sector-based searching mechanism is always valid as terminals in the remaining sectors are obviously far away from the terminal in consideration.

The computation of the exhaustive search is in general of order $O(n^2)$ complexity. However, the computation time is reduced considerably if the decision is made on top of the sector based terminal identification.

Note that the selection mechanism is local to the cell, i.e, when a relay terminal is selected, the selection is only within the cell and the impact of other transmitting terminals outside of that cell is not included. This is due to the fact that the AP has little knowledge about the ongoing transmissions outside its cell and hence the effect of co-channel interference cannot be predetermined. Nevertheless,

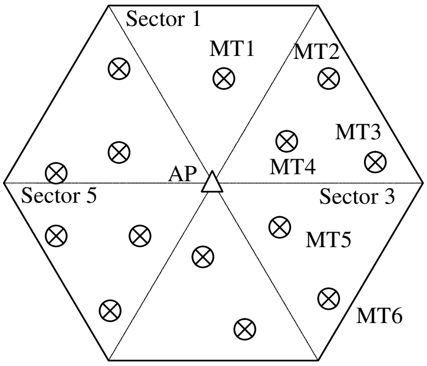


Figure 6.6: An example cell where terminals are identified by a sector

this local selection scheme serves its purpose considerably. Later when terminals sense the presence of interference, they readjust their data rate by first recalculating their actual PER.

6.3.2 Relaying Modes

As described earlier, terminals switch to relaying mode as long as there is a terminal which give a higher effective data rate than the direct transmission for the same given time slot and traffic size. Thus, in one communication frame, both directly communicating and relaying terminals can be scheduled. The relaying can further be categorized in three cases.

One-frequency Relaying

Every cell is assigned a single frequency which is responsible for the entire communication (both direct and relay communications) within the cell according to a well defined frequency reuse pattern such as 3, 7 or 19 frequency reuse.

The communication impediment in this case is mainly due to the interference coming from the neighboring cells that use the same frequency. The intensity of the interference varies depending on the frequency reuse pattern. If a small number frequencies is reused (for e.g., 3 frequency reuse pattern), the same frequency appears repeatedly in other cells which are not too far from the cell in consideration and considerably affects the ongoing transmission. Conversely, if the number of frequency to be reused is large enough (e.g., 19 frequency reuse pattern), the frequency reuse distance becomes larger and the effect of interference is relatively lower.

In direct communication, there is only a single sender-receiver pair at a time as the AP is always involved in the communication. Whereas in multi-hop communication, there are more sender-receiver pairs. It is possible to scheduling these pairs as concurrent transmissions within the cell on a single frequency so that interference will not be high. However, for relatively small sized cells, as in the case considered here, it is hardly feasible to schedule such concurrent transmissions due to the then ensuing increased interference and packet error rate. In order to make use of the relaying advantage, a second frequency can be added from the neighboring cell which will then be used for relaying to combat the increased interference.

Two-frequency Relaying

Due to the scares frequency resource, it is not always easy to dedicate a separate frequency within a cell to support relaying. Additional frequency is only available by “borrowing” or “recycling” the neighboring cell’s frequency. The available sender-receiver pairs due to multi-hop communication can now be scheduled concurrently within the cell in two different frequencies.

In this relaying mode, two frequency reuse patterns are distinguished: By *macro frequency*, it refers to the standard, static frequency reuse pattern assigned to every AP in the network. The *micro frequency* reuse pattern is the one which determines how frequencies are recycled close to the AP for the sake of relaying. An example of such frequency recycling is shown in Figure 6.7. This second, recycled frequency is used only in the interior of a cell, close to the access point. Consequently, the interference from the lending cell is minimized. However, there is still a trade-off between the increased interference that the recycled frequency will cause on the overall system and the capacity gains within the cell achieved through two-frequency relaying.

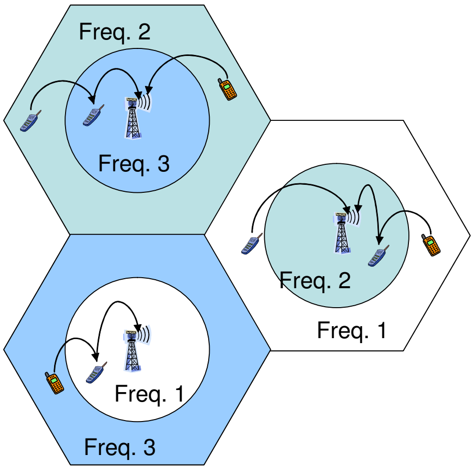


Figure 6.7: An example frequency recylcing

One important condition for two-frequency relaying is the existence of independent communication entities. An entity cannot send and at the same time receive in different frequencies. If sender-receiver pairs do not fulfill this condition, they are not valid candidates to be scheduled in the two-frequency frame slot.

Sectored Frequency Relaying

Similar to the two-frequency case, recycled frequencies are used to schedule concurrent transmissions. The difference is, instead of having only two-frequencies, the cell will have the macro frequency which is used for the primary communication and a number of recycled frequencies each in the six sectors of the hexagonal cell. The frequency assignment depends on the frequency reuse pattern used. An example of such sectored frequency relaying is shown in Figure 6.8. As in the previous case, concurrent transmissions are scheduled as long as there are independent communication entities. Otherwise, the relaying would remain as in one-frequency relaying case.

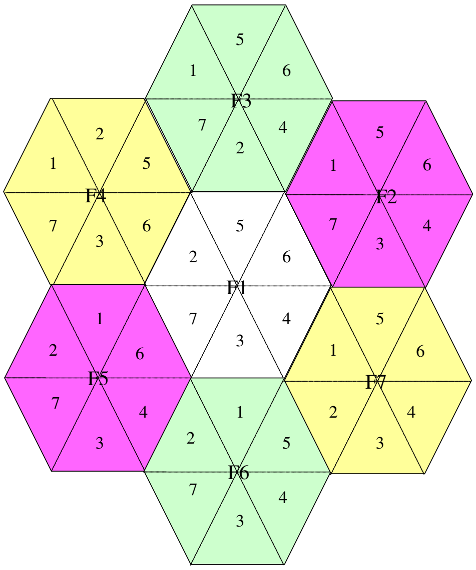


Figure 6.8: An example of sectored frequency recycling for 7-frequency reuse pattern

6.4 Fairness Requirement

In a TDMA-based cellular wireless system like HiperLAN/2, mobile terminals need to send their traffic to the desired destination in a given time slot. When attempting to maximize the capacity of such a cell, the best way is to assign all time slots to the terminals that are closest to the AP as it will realize the highest throughput. Evidently it is not intended, in this work, to maximize the capacity by being unfair to other transmitting terminals which are still waiting for resource allocation. It is also assumed that all terminals have homogenous requirements. Hence, some fairness requirements need to be demanded.

Scheduling terminals according to some fairness condition is the next step after the relaying decision [82]. Two fairness schemes are identified in this thesis: the "uniform slot size" and "uniform traffic size" schemes. In the "uniform slot size" scheme, terminals obtain an equal share of the total frame time when scheduling the communication. The “uniform traffic size” scheme is used when all terminals are allowed to send a uniform amount of data in slots of varying length depending on their modulation and, ultimately, on their distance from the access point. These two options reflect different perspectives on how to share system resources – the first one might be more provider oriented, the second one is rather customer oriented (“I don’t care where I am, I want the same effective data rate as everybody else”).

In addition, when terminals have a variable bit rate load with a randomly sized traffic, “demand oriented” fairness scheme is identified. In this scheme, terminals share the TDMA frame according to the traffic size they have, i.e., terminals with large traffic get a larger share and those with small traffic get less. But to guarantee a minimum frame time slot and to avoid unfair use of the time resource by few terminals with very large traffic size, only half of the TDMA frame is shared according to the traffic size. The rest half is equally distributed among the terminals. Similar to the “uniform traffic size” scheme, the “demand oriented” is customer oriented as well.

The total capacity of a cell which uses a “uniform slot size” scheme is simply the arithmetic mean of the terminal’s data rates (whether it is derived by relaying or not is not relevant), times the number of terminals. For the “uniform traffic” scheme, this is more complicated. Looking at the example from Section 6.7 gives an idea: To transmit, say, 1 Mbit from each terminal takes $1/6$, $1/36$, and $1/27$ seconds, respectively. Hence, 3 MBit are transported in $1/6 + 1/36 + 1/27$ seconds, yielding a throughput of $3/(1/6 + 1/36 + 1/27)$. More generally, the uniform traffic scheme corresponds to the *harmonic* mean of the individual data rates. It is worthwhile to point out that the harmonic mean corresponds to the behavior of IEEE 802.11: When every terminal gets the same chance to access the channel to transmit data packets of the same size, traffic fairness ensues. The “performance anomalies” of IEEE 802.11 [48] are, essentially, only the result of the harmonic mean’s sensitivity to outliers. Relaying should be particularly beneficial in this fairness scheme as it benefits in particular far terminals, which severely affect the cell capacity under a harmonic mean averaging process.

Consider the example again. For the “uniform slot” fairness scheme, relaying improves the cell capacity from (6+36+27)/3 = 23 MBit/s to (12+36+27)/3 = 25 MBit/s (a gain of about 8.6 %); under a “uniform traffic” scheme, it is improved from 3/ (1/6 + 1/36 + 1/27) = 12.96 MBit/s to 3/ (1/12 + 1/36 + 1/27) = 20.25 MBit/s, a gain of about 56 %!

Figure 6.9 summarizes the different fairness schemes for direct, one- and two-frequency relaying cases. The example is too small to show the effect of keeping the second frequency close to the access point, nonetheless the overall picture should be clear.

6.5 Scheduling

Routing decision alone is not enough when optimizing the capacity of cellular networks. Even if there are a number of concurrent sender-receiver pairs due to relaying, if they are not properly scheduled in the TDMA communication frame, these pairs could be another source of interference to the system,

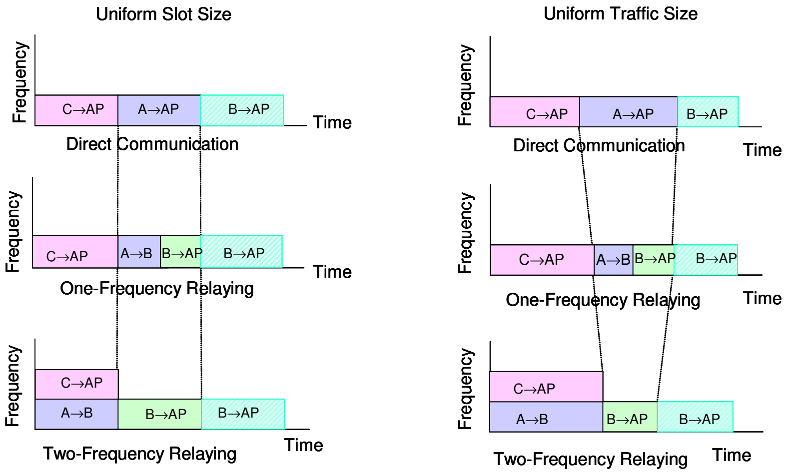


Figure 6.9: Time-freqency representation of the two fairness schemes, with hypothetic terminals A, B and C; A relaying via B

which could in turn affect the overall performance of the system. Thus, scheduling plays a crucial role in providing an integrated improvement in the system capacity.

As indicted earlier, the AP is responsible in scheduling and setting up the TDMA communication frame. Based on the knowledge that AP has, it calculates the time slot every terminal should be alloted. The scheduling depends on the fairness conditions and communication modes discussed in Section 6.3.2 and 6.4. The scheduling mechanism are presented as follows:

1. **Direct communication with uniform slot-size fairness**

   This condition is relatively straightforward. If there are n actively transmitting mobile terminals, each will have $\frac{1}{n}$ of the time frame. Each terminal then send its traffic to the AP in a sequential manner. The amount of traffic to be sent depends on the distance the terminal is away from the AP. Let r_j^i be the effective data rate used by terminal j in cell C_i. The total throughput at the AP is:

$$\tau = \sum_{j=1}^{N} (1 - \mathrm{PER}_{\mathrm{j}}^{\mathrm{i}}) \cdot r_j^i \qquad (6.6)$$

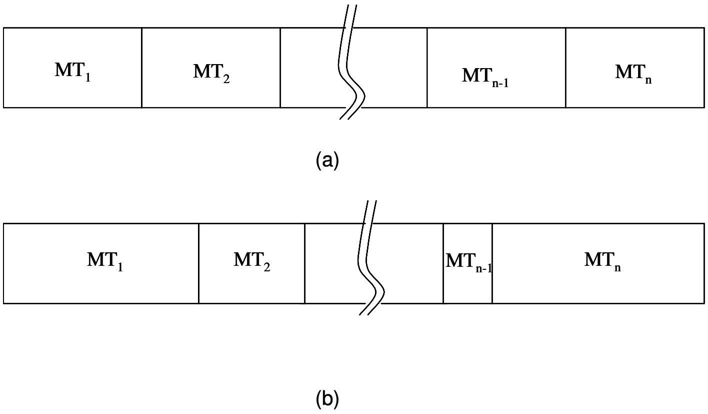
2. **Direct communication with uniform traffic-size fairness**

   Every terminal transmits the same amount of traffic for the allocated time slot. But the size of the slot varies according to the data rate each terminal is sending. Closer terminals to the AP use faster data rate and hence their time slot is smaller. The amount of traffic is determined as follows. Let r_j^i be a non-zero effective data rate used by terminal j in cell C_i. If each terminal is allowed to transmit ξ traffic size, the total throughput at the AP is:

$$\tau = \sum_{j=1}^{N} (1 - \mathrm{PER}_{\mathrm{j}}^{\mathrm{i}}) \cdot \xi \tag{6.7}$$

where ξ is given by:

$$\xi = \sum_{j=1}^{N} \frac{1}{r_j^i} \tag{6.8}$$



(a)

(b)

Figure 6.10: TDMA frame with (a) uniform slot-size and (b) uniform traffic size

3. **Direct communication with demand oriented fairness**

   Similar to the uniform slot-size case the n terminals equally share half of the frame. The other half is shared according to the traffic size of the terminals. Let x_i be the traffic size of terminal i. The slot size of each terminal is then $\left(\frac{1}{N} + \frac{x_i}{\sum_{i=0}^{N} x_i}\right)$. According to the allotted time, the terminals sequentially send their traffic.

4. **One-frequency relaying with uniform slot-size fairness**

   Again here each source terminal shares the communication frame equally. If relaying is beneficial for a terminal, it splits its own time slot into “relayed to relay” and “relay to AP” slots so that the fairness among other terminals kept unaffected. Similar to the direct case, the transmission is sequential.

5. **Two-frequency relaying**

   Two overlapping TDMA frames are needed to schedule the transmissions with the macro frequency of the cell and with the micro frequency. The second frequency is used in the vicinity of the AP for relaying purpose. To assign active terminals for simultaneous transmission, the location information and interference situation are relevant. The following assignment rules are also required to schedule the simultaneous transmission:

Rule 1

Mobile terminals cannot be assigned in the primary and overlapping slot simultaneously.

Rule 2

If a potential two-frequency transmitter does not satisfy rule 1, it degenerates to one-frequency relaying and sends its traffic by splitting its own slot.

If a terminal serves as a relay for three or more other terminals, it is less likely that the terminals satisfy Rule 1. This is specially true for cells with small number of terminals.

6.6 Frequency Recycling Pattern

Two-frequency relaying is already suggested as a means to further improve the capacity achieved through relaying. Due to the spectrum resource scarcity, the second frequency is obtained from the neighboring cells. The number of frequencies to be recycled depends on the type of frequency reuse pattern the cellular network is using. From the typical patterns used here, for e.g., 3 frequencies will be recycled when the reuse frequency is 3 and the same is true for the 7 and 19 frequency reuse patterns.

In the cases when there are plenty of frequencies available, for e.g., as in the case of 19, an alternate frequency pattern is used for the micro frequency distribution. Instead of using all the available channel, part of it used for macro and another part is dedicated for the micro frequency reuse pattern. This dedicated frequencies are used for relaying purpose similar to the other recycled ones. One particular example is the use of 7 frequencies each for both macro and micro frequency patterns and 5 more left free. This is expected to reduce the additional interference in the entire system due to simultaneous transmissions. Another interesting issue in relation to dedicating frequencies from the already available frequency spectrum is dynamically splitting the entire bandwidth into a number of channels and use these channels separately for the primary communication and multi-hop communication purpose. These requires factors such as prior knowledge of the system load situation, the frequency resource management and flexibility of the system hardware. However, this is not the scope of the thesis and left for further study.

6.7 Example Scenario

Let us consider an example scenario which can summerize the mechanism discussed so far in this chapter. The scenario consists of a simple cellular network with three terminals A, B, and C as shown in Figure 6.11; assume that B is a potential relay for A and C is communicating only directly with the AP. Assume also that the medium is shared equally between A, B, and C, resulting in a net data rate of 2, 12, and 9 MBit/s for each terminal, respectively.

Evidently, A can improve its direct rate from 2 MBit/s to at least 3 (=18/2/3) MBit/s if the time slot for the A to AP communication is split in two equal parts, used for communicating from A to B

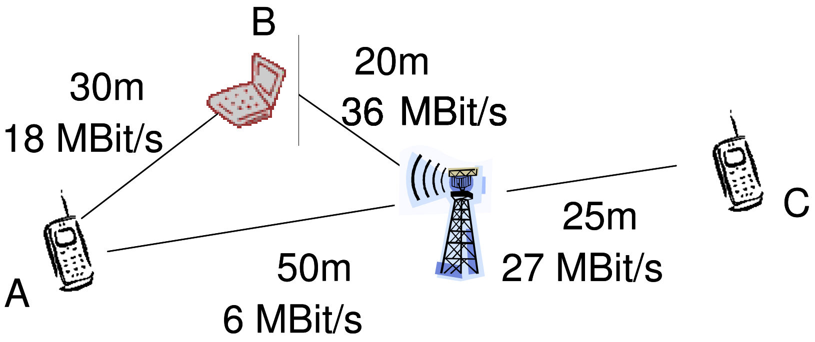


Figure 6.11: A simple cellular network scenario

at 18 MBit/s and from B to AP at 36 MBit/s (the division by two accounts for the need to transmit A's packet twice, the division by three models the sharing of the medium with B's and C's traffic). Splitting the time slot into sub-slots of length 2:1 between the A–B and the B–AP links results in a net data rate for A to the AP of 4 = (18*2/3)/3 = (36*1/3)/3 MBit/s. This *rate-adaptive relaying* has doubled A's net data rate.

Moreover, when using relaying, the AP is idle when A transmits to B. This time can be used for C's transmission to the AP. To schedule these transmissions concurrently, a second frequency is necessary (doing so in the same frequency is possible, in very large cells, but in general it is not advantageous because of the increased intra-cell interference [37]).

In summary, the following are key issues addressing in improving the capacity of the wireless cellular networks.

- a transmission power and rate adaptation problem for each individual link
- a routing problem to select, for each terminal, a relaying terminal or to decide not to relay at all
- a scheduling problem how to assign time slots based on the individual links' data rates and on the option of a second frequency, and
- proper frequency recycling patterns

In the next chapters, the simulation tool used to implement the mechanisms proposed is discussed and the respective simulation results for different scenarios is followed.

Chapter 7

The Simulation Tool

7.1 Introduction

In Chapter 5, the potential benefits of relaying in wireless cellular networks is mathematically analyzed. For the analysis, separate analytical models are used to independently study the factors that influence the performance of the network–the transmission power, noise and interference. But the performance of a realistic network heavily depends on the combined effects of different factors. Analytically studying the network by parametrically representing all the factors is rather cumbersome and it is difficult to come up with a unified solution. Therefore, it is apt to use a simulation model that meets expectations. This chapter is dedicated to discuss the simulation tool that is essential to implement the proposed mechanisms in Chapter 6. In the next sections the simulator architecture and its entities are described.

7.2 Simulator Overview

Basically the simulator is designed to flexibly represent the different aspects of a multi-hop cellular network which was otherwise difficult to mathematically represent. It is written in C^{++} and provides different functionalities that allows to evaluate the performance of the wireless network.

The architecture of the simulator is shown in Figure 7.1. It comprises of five basic blocks: the channel model, the physical layer, the input, the simulator engine and the evaluation. The *channel model* specifies the propagation model of the wireless system. It also provides a functionality to calculate the interference pertaining to the system. The *physical layer* stipulates the radio map [31] that helps to estimate the channel gain. The transmission power as well as the modulation to be used are specified in this block. Off-line generated inputs, simulation parameters and other simulation options are furnished to the simulation engine by the *input* block. The core of the simulator in the *simulation engine* which actually simulates the network performance based on the physical layer and

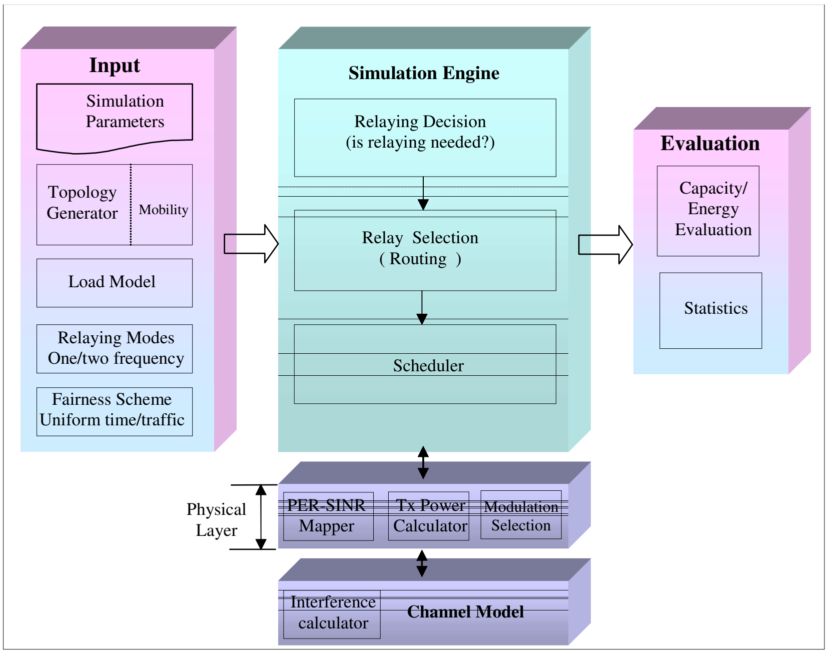


Figure 7.1: Simulator architecture

channel model specifications and the inputs provided. The simulation output is then analyzed in the *evaluation* block. The architecture itself lends to simple modifications and to introduction of new functionalities. The next sections shed light on the basic blocks of the simulator architecture.

7.3 Input Block

In this section, the different components of the input block shown in Figure 7.1 are presented.

7.3.1 Topology Generator

Topology generation is the basis for analysis and evaluation of multihop cellular network. It is responsible for generating cell clusters and mobile terminals that typify the network. Depending on the number of available frequencies to be reused, the topology generator is also responsible to make static macro- and micro-frequency reuse assignment to each cell. Moreover, whenever there is mobility in the network, the mobility trace files are generated here. Ultimately the off-line generated topologies are passed to the simulation engine as an input. The cell cluster and terminal generation as well as the terminal mobility are described as follows:

Cell Cluster Generation

The initial step in topology generation is to create a cluster of cells which are assumed to cover a certain network service area. The cells have regular hexagonal shape as discussed in Section 3.3.1 and each cell is divided into six sectors. The cells are identified by the name and position of their AP which is located at the center of each cell, i.e., by AP_{name}, AP_x and AP_y. And the sectors in each cell are simply numbered from one to six and the corresponding number is used to identify the sectors.

For a given cell radius, the first cell is generated at the origin of a certain service area. The first cell is the *reference cell* for the performance evaluation of the entire network. Then, the rest of the cluster is generated around this reference cell. The surrounding cells are chiefly needed for two main reasons: they provide interference for the reference cell and they avoid edge-effects during simulation. The cluster cells are generated as follows: Let γ be the radius of the inscribed circle in the hexagonal cell, i.e., γ is the apothem of the cell radius. Six cells that surround the reference cell are generated to form the 7-cell cluster at centers given by:

$$\begin{aligned} x &= 2\gamma \cdot \cos\left(\frac{\pi}{6}(2k-1)\right) && \text{and} \\ y &= 2\gamma \cdot \sin\left(\frac{\pi}{6}(2k-1)\right) && \text{for } k \in \{1, ..., 6\} \end{aligned} \tag{7.1}$$

Once the cluster is formed, the generation of additional cluster cells depend on the frequency reuse pattern chosen. For example, for 19-frequency reuse pattern, 12 more cells need to be generated that surround the 7-cell cluster, whose centers are located at:

$$x = \begin{cases} 4\gamma \cdot \cos\left(\frac{\pi}{6}(2k-1)\right) & \text{for } k \in \{1, ..., 6\} \\ 2\sqrt{3}\gamma \cdot \cos\left(\frac{\pi}{3}(m-1)\right) & \text{for } m \in \{1, ..., 6\} \quad \text{and} \end{cases} \tag{7.2}$$

$$y = \begin{cases} 4\gamma \cdot \sin\left(\frac{\pi}{6}(2k-1)\right) & \text{for } k \in \{1, ..., 6\} \\ 2\sqrt{3}\gamma \cdot \sin\left(\frac{\pi}{3}(m-1)\right) & \text{for } m \in \{1, ..., 6\} \end{cases} \tag{7.3}$$

Afterwards the cluster patterns are repeated at reuse distance of $D = \sqrt{3}RN$, where R is cell radius and N is the number of frequencies to be reused, as discussed in Section 2.6. An examples of such cell clusters are shown in Figures 7.2 and 7.3.

Eventually, the macro- and micro-frequency reuse assignment is done for each cell. For sector-based relaying mode, the sectors in a cell are also assigned a separate frequency. Note that the micro- and sector-frequencies are recycled frequencies from the neighboring cells.

Terminal Generation

Terminals in the cellular network are generated per cell basis depending on the number of terminals specified in the simulation parameter. These terminals are uniformly distributed in the entire simu-

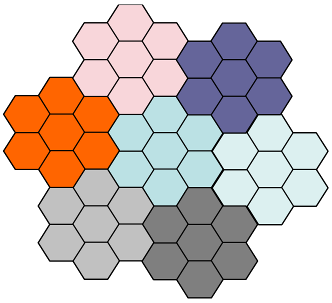


Figure 7.2: Cell layout with 7-frequency reuse pattern

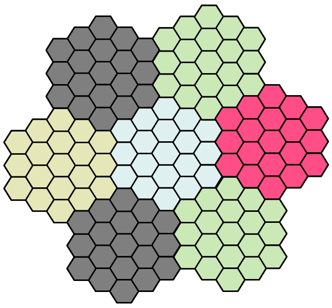


Figure 7.3: Cell layout with 19- frequency reuse pattern

lation area and are identified by their MAC-ID, position MT_x and MT_y and the sector number they belong. They also have different attributes to identify if they are directly communicating with their AP or if they are selected as a relaying terminal for others or if they are sending their traffic via a relaying terminal.

Mobility

If mobility is selected, terminals randomly choose their speed and direction from $[v_{min}, v_{max}]$ and $[0, 2\pi]$, respectively and move for a specified number of *steps* [15]. Steps in this context are the number of simulation cycles the terminals continue to move before choosing a new speed and direction. The topology generator generates a trace file for each step. If terminals detect cell boundary, they will bounce and continue to move with a speed chosen prior to the bounce.

7.3.2 Load Model

The load model generates a constant bit rate (CBR) for the transmitting terminals. It also specifies the source and destination of the traffic it generates. It is assumed that there is an infinite source of traffic and the network is always performing at full load. In addition, the model generates a variable rate load with a random traffic size with random intervals.

7.3.3 Simulation Parameters and other Inputs

This module specifies the parameters to be used in the simulation. These are: number of cells, number of terminals in a cell, Noise level, target PER, pathloss coefficient and number of frequencies available for reuse.

The other inputs which are required by the simulator engine are the relaying mode and fairness scheme selection modules. In the former case, direct, one-frequency, two-frequency and sector-based relaying options are given. In the latter case, a uniform time slot or uniform traffic size fairness options are given to share the system resource.

7.4 Channel Model

In infrastructure based cellular network, the AP is the central controller of every ongoing transmission. Before each terminal starts sending its traffic according to the TDMA frame schedule, the AP should estimate the channel gain of each terminal. One way of determining the channel gain is using the concept similar to HiperLAN/2's "radio map" [30], i.e., the AP knows the received signal strength (RSS) of each terminal and which mobile terminal is in radio contact with which other mobile terminal. This module provides the channel gain estimation of the "radio map". For the estimation, it needs location information of the terminals and the propagation model of the communication channel described in Section 3.2. Furthermore, the module is responsible for calculating the interference level for each destination terminal.

Interference Calculator

When dealing with multi-cellular network where there is extensive use of frequency reuse, the existence of interference is inevitable. This module is responsible to calculate the interference level for each destination terminal based on the channel gain. The source of interference varies according to the type of communication mode selected: for direct communication and one-frequency relaying modes, the co-channel interference is due to the macro frequency reuse whereas for two-frequency and sector-based relaying the interference is due to both macro- and micro-frequency reuses.

The interference calculation is solely done for the reference cell and all other transmissions outside of the reference cell are taken as sources of interference for the reference cell. Assume that the reference cell AP has the information about the ongoing transmission in other cells, including the location information of transmitting terminals, the interference is calculated as follows: Let terminal i is sending to j in the reference cell. Assuming that another terminal k is simultaneously transmitting in a co-channel cell at a transmission power of P_{tx}^{k}. The interference $P_{I,j}^{k}$ at the receiving terminal j due to k in the reference cell is:

$$P_{I,j}^{k} = \frac{P_{tx}^{k}}{10^{\frac{1}{10}(46+10\alpha\log d)}} \tag{7.4}$$

where α is the pathloss coefficient and d is the distance between terminal j and k. If there are other simultaneous transmissions in other co-channel cells, the interference at the receiving terminal j is due to the aggregation of the transmissions. The calculated interference is needed to determine the signal-to-interference-noise ratio which is eventually used to adaptively choose the transmission power and rate of communicating terminals.

7.5 The Physical Layer Block

This block sets the transmission power and the modulation rate which are needed both by the scheduler in the simulation engine and by the channel model. In the next subsections, the physical layer modules are presented.

7.5.1 PER Mapping

The PER-SINR mapping module is responsible to map the PER for the signal-to-noise/interference ratio of a radio map using the exponential curve-fitting discussed in Section 5.2.1. Figure 7.4 shows the curve fitting of the PER to C/I for all physical layer modes of HiperLAN/2 in [61]. It is also responsible to determine if the received signal at a destination terminal is below the receiver sensitivity. The sensitivity levels for each modulation are shown in Table 3.2.

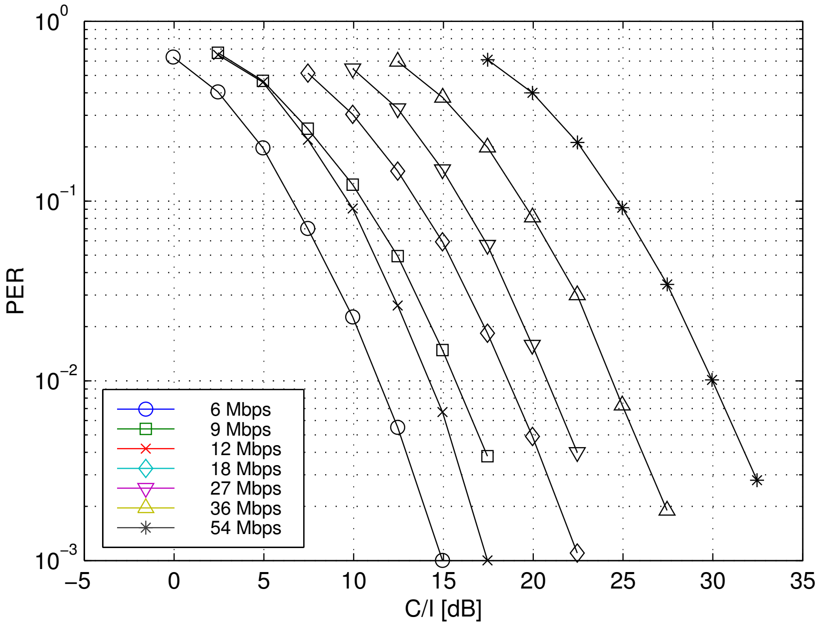


Figure 7.4: An exponential curve fitting for all physical layer modes

7.5.2 Modulation Selection

In this module, the threshold SINR that satisfies the target PER is first calculated for each of the seven modulations from the PER-SINR mapping. Next, the maximum possible distance that can be covered with the threshold SINR, maximum possible transmission power and a given pathloss coefficient is calculated for the seven modulations. Then, set of curves similar to Figure 6.1 are generated in the end. Modulation selection is finally done on the basis of these set of curves and distance between the two communicating terminals.

7.5.3 Transmission Power Calculation

For a selected modulation, the threshold SINR that satisfies the target PER is first obtained from the PER-SINR mapping. Then transmission power need to cover a given communication distance d is calculated as:

$$P_{tx}[dB] = \mathrm{SINR}_{th} + \mathrm{N} + 46 + 10\alpha \log d \tag{7.5}$$

when interference level increases, the actual PER will also be high. But to maintain the target PER, modulation rate is reduced step by step and the transmission power is accordingly recalculated. Thus the transmission power and modulation are adaptive to the interference level.

7.6 Simulation Engine

The simulation engine is the main part where the simulation is running. It takes the load model, the corresponding topology and other parameters as an input. By making use of the radio map information, the AP in the reference cell first makes a decision if relaying is necessary. Afterwards the relay selection and scheduling is done depending on the relaying mode and fairness scheme choices from the input block. The implementation of this block is as discussed in Chapter 6.

Ultimately, the output of the simulation engine is a TDMA frame schedule. This schedule stipulates the slot allocation of every transmitting terminal together with the corresponding transmission power and rate. The resource request of the transmitting terminals and other administrative broadcasts that the AP performs are not included in this simulation.

7.7 Evaluation Block

From the resulting frame schedule, the total throughput GP is calculated for the reference cell depending on the relaying modes and fairness schemes.

Direct communication with uniform slot-size fairness:

$$\mathrm{GP}_{dir} = \sum_{i=1}^{N} (1 - \mathrm{PER}_i) \cdot \mathrm{R}_i \tag{7.6}$$

One-frequency relaying with uniform slot-size fairness:

$$\mathrm{GP}_{1F} = \sum_{i \in S} (1 - \mathrm{PER}_i) \cdot \mathrm{R}_i + \sum_{j,k \in N-S} (1 - \mathrm{PER}_j)(1 - \mathrm{PER}_k) . (\frac{1}{\mathrm{R}_j} + \frac{1}{\mathrm{R}_k}) \tag{7.7}$$

where S is the set of terminals in frame that direct communication and the rest N-S terminals are scheduled for one-frequency relaying; j is a terminal that selects k as its relaying terminal and R is the corresponding data rate used by the terminals.

Two-frequency relaying with uniform slot-size fairness:

$$\mathrm{GP}_{2F} = \sum_{i \in S} (1 - \mathrm{PER}_i) \cdot \mathrm{R}_i + \sum_{j,k \in N-S-P} (1 - \mathrm{PER}_j)(1 - \mathrm{PER}_k).(\frac{1}{\mathrm{R}_j} + \frac{1}{\mathrm{R}_k}) + \sum_{l \in P} (1 - \mathrm{PER}_l) \cdot \mathrm{R}_l \tag{7.8}$$

where P is the set of terminals in the frame that are scheduled for two-frequency relaying and the N-P-S terminals use one-frequency relaying. Note that even in two-frequency mode, terminals can be scheduled to send in one-frequency if they cannot be scheduled in the second-frequency frame as discussed in Section 6.5.

With uniform traffic-size fairness

Let the traffic size and effective data rate of each terminal be ξ and r_i, respectively. The resulting total throughput at the AP is:

$$GP = \sum_{i=0}^{N} (1 - \mathrm{PER}_i) \cdot \xi \tag{7.9}$$

where ξ is given by:

$$\xi = \sum_{i=1}^{N} \frac{1}{R_i} \tag{7.10}$$

With demand oriented fairness

The throughput computation is similar to the uniform slot size fairness except for the slot size factor:

$$\mathrm{GP}_{dir} = \sum_{i=0}^{N} \frac{1}{2} \left(\frac{1}{N} + \frac{x_i}{\sum_{i=0}^{N} x_i} \right) \cdot (1 - \mathrm{PER_i}) R_i \tag{7.11}$$

where x_i is the traffic size of terminal i.

Finally the total average throughput is computed from the different inputs and topologies provided to the simulation engine. A confidence interval at 95% confidence level is also computed for the resulting average throughput.

7.8 Summary

The simulator is basically used to implement the integrated solution proposed to improve the capacity of a cellular network using relay communication. Further assumptions used in the simulator are summarized as follows: When relaying is selected, a maximum of two hop or single-relay-hop is used for the reasons discussed in Section 6.3. The choice is reasonable since the cell sizes considered are relatively small. Besides, the frequency switching and relay management overhead that may be required to support multi-hope is reduced. To avoid an unbalanced traffic assignment to relaying terminals which are close to the AP, a single relaying terminal is allowed to relay only for up to five other terminals.

It is assumed that there are up to 19 non-overlapping channels available in the 5 GHz communication range. The available channels are to be reused in 3, 7 and 19 frequency reuse patterns. It is also assumed that there is sufficient suppression of spectral overlap and hence adjacent channel interference is safely ignored in the system.

As indicated earlier, the PER to SINR mapping used in the simulation is obtained from the performance of representative HiperLAN/2 system in [61]. In the mapping a log-normal fading with a standard deviation of 2dB is added to model shadowing. Therefore, no additional parameter for fading is considered in this simulation. A noise level of -90 dBm is used in the simulation. Furthermore, a symmetrical pathloss is assumed and hence the treatment of up-link and down-link communication is similar. The simulation is done only at MAC level and no error coding mechanism is included as well.

Lastly, the essential parts of the multi-hop cellular network simulator used is described in this chapter. Although the assumptions taken in the simulator rely on HiperLAN/2 system, the implementations are independent of the link technology. Thus, the analysis and results can be extended to IEEE 802.11a system without rigorous modifications.

Chapter 8

Simulation Results

8.1 Introduction

The analytical results in Chapter 5 substantiate the potential benefit of relaying to improve the capacity of cellular wireless networks. The results indicted are of limited scope as the analysis include only few but essential factors that influence the performance of a wireless network. However, it is vital to take the possible factors into consideration when studying such networks. But since mathematically representing all factors is difficult, simulation based evaluation is opted to study the network performance. The mechanism proposed to improve the capacity of a cellular network are implemented in the simulator described in the previous chapter. The implementation is first done for a single cell network. Later, it is extended to a multi-cells network. This chapter is thus devoted to discuss the evaluation of the proposed mechanism for both single and multi-cell networks.

8.2 Single Cell Network

The single cell simulation is done under the assumption that the frequency reuse distance of the multi-cellular wireless network is large enough to safely ignore the co-channel interference in the system. Hence, pathloss and noise are the factors that affect the received signal as in transmission power limited system.

Apart from the direct communication, two relaying schemes are considered here– one- and two-frequency relaying. When one-frequency relaying is used, both the terminal to relay and relay to AP are scheduled within the same slot which would otherwise be used for direct communication between the terminal and the AP, i.e., the relay terminal does not require an extra slot in the MAC fame. In two-frequency relaying case, the second frequency is from a “neighboring cell” and is used close to the AP only for relaying purpose. If there is any simultaneous transmission, they are scheduled in a separate TDMA communication frame using the second frequency. The effect of additional interference in the system due to the second frequency is not treated in the single cell simulation. Fairness among terminals is maintained by either scheduling the terminals to get an equal share of the total frame time or allowing all terminals to send a constant amount of data in slots of varying length depending on their modulation [81].

In the simulation, 55 different random placements of terminals on a square area of 70 m x 70 m are used, the access point being in the middle. A maximum target PER of 1%, 3%, 5% and 10% and different pathloss coefficients are used in the throughput computation. Average throughput at the AP is obtained by averaging the total throughput from all the terminals over the 55 random placements for a given target PER and pathloss coefficient α. The confidence interval for the results is computed at 95% confidence level.

The resulting average throughputs for 20 terminals in the cell at varying pathloss coefficients and a target PER of 1% are shown in Figure 8.1 and 8.2 for uniform slot size and uniform traffic fairness schemes, respectively. For small α, both schemes are capable of fully utilizing the access point's maximum throughput. It is possible to attain the maximum throughput only using direct communication and hence there is no need for the terminals to switch to relaying mode in the case of small α. But as α increases, the range of communication over which the target PER condition can be met at a fixed modulation and limited transmission power decreases. Hence, the attained throughput at the access point also decreases. As shown, relaying becomes essential for higher values of α. The average throughput beyond $\alpha = 3.2$ is not shown since direct communication between the terminals and the AP is not always guaranteed within the cell for maximum target PER to be maintained.

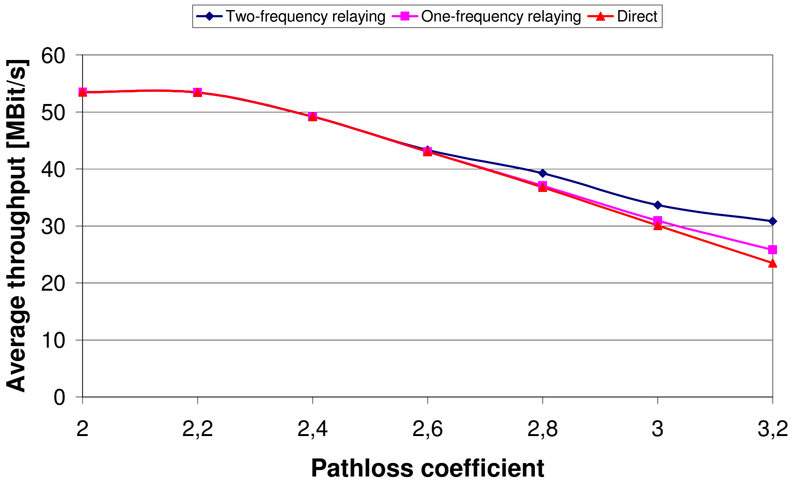


Figure 8.1: Average throughput for single cell as a function of α for a *uniform slot size* fairness scheme and a target PER of 1%

Figure 8.3 and 8.4 show the average throughputs in two-frequency relaying mode for both fairness schemes using different target PERs. The achievable throughput at the AP for larger target PERs is

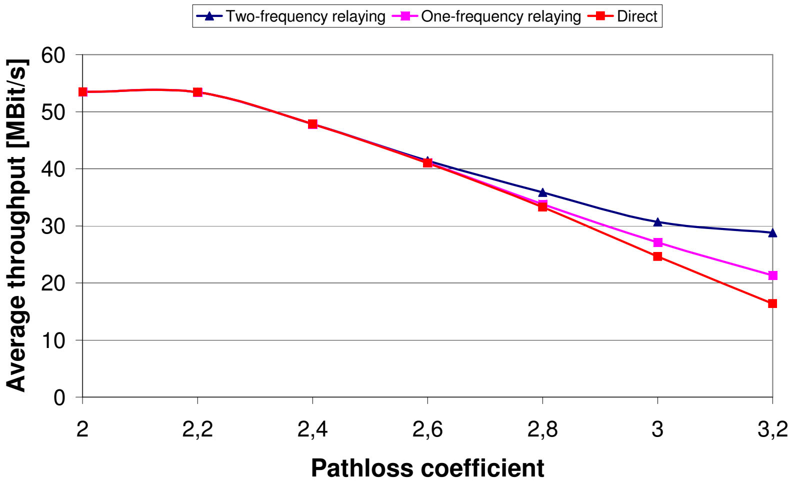


Figure 8.2: Average throughput for single cell as a function of α for a *uniform traffic* fairness scheme and a target PER of 1%

relatively higher than that of smaller PERs. However the target PERs cannot simply be made larger as there will not be communication due to high packet loss.
The average throughputs at the AP for $\alpha = 3.2$ and different number of terminals within the cell for both fairness schemes are shown in Figure 8.5 and 8.6. In both cases relaying performs always better than direct communication. Figure 8.7 and 8.8 show the average throughput for two-frequency relaying mode only in both fairness schemes at $\alpha = 3.2$ using different target PERs.
As the number of terminals increase, there is a considerable increase in the average throughput at the AP. In extreme examples, two-frequency relaying almost doubles throughput achieved by direct communication. Even using only one-frequency, the throughput achieved by the direct communication is improved by about 40%. The dependency on the number of terminals is essentially a stochastic effect: the far terminals have a higher probability of finding a relay terminal, more often enabling faster modulations. Hence, relaying does generate more capacity when it is most sourly needed.

For all fairness schemes, relaying out performs direct communication. But compared to the uniform traffic fairness scheme, the uniform time slot fairness scheme shows a better performance. Selecting appropriate resource sharing scheme is, thus, equally important when improving the network performance.

8.3 Multi-Cell Network

In the previous section, it is shown how relaying improves the throughput at the AP of a single cell. The impact of interference was not included in the single cell treatment case by assuming very large frequency reuse distance to ignore co-channel interference. Now in this section a number of cells are

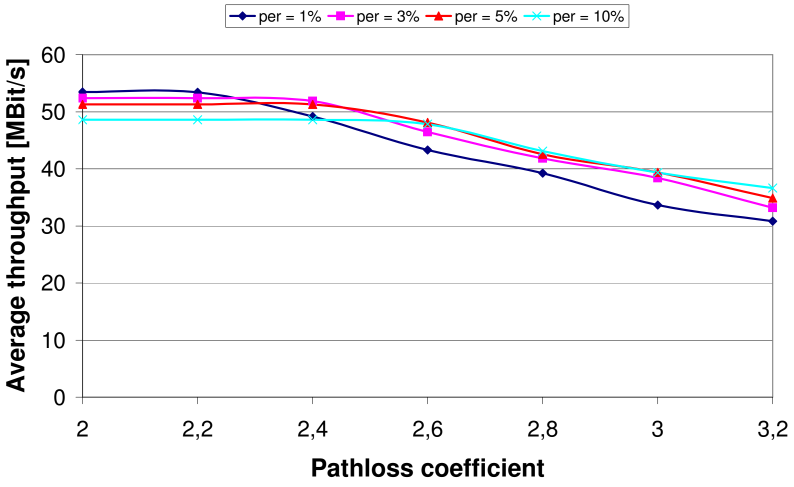


Figure 8.3: Average throughput for single cell as a function of α, for a *uniform slot size* fairness scheme and two-frequency relaying, for target PERs of 1%, 3%, 5% and 10%

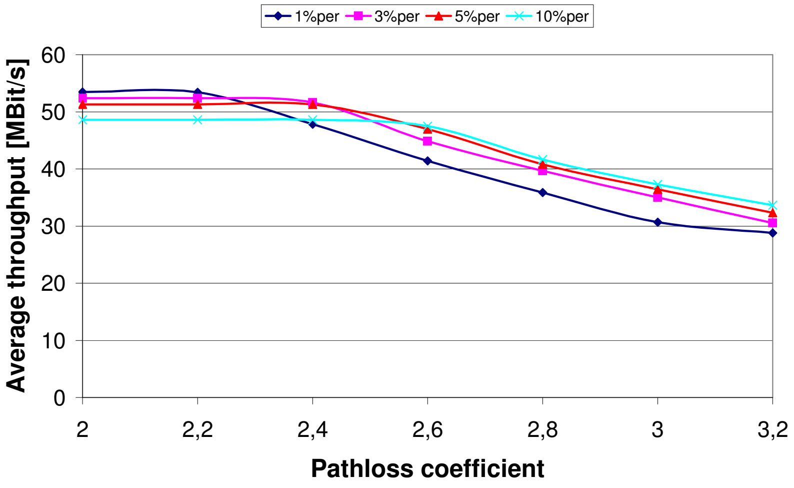


Figure 8.4: Average throughput for single cell as a function of α, for a *uniform traffic* fairness scheme and two-frequency relaying, for target PERs of 1%, 3%, 5% and 10%

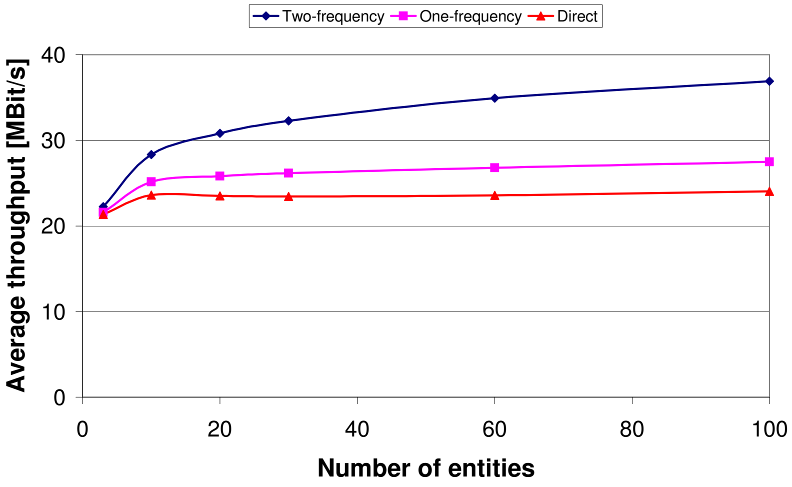


Figure 8.5: Average throughput for single cell as a function of terminals in cell for a uniform slot size fairness scheme, pathloss coefficient $\alpha = 3.2$ and target PER of 1%

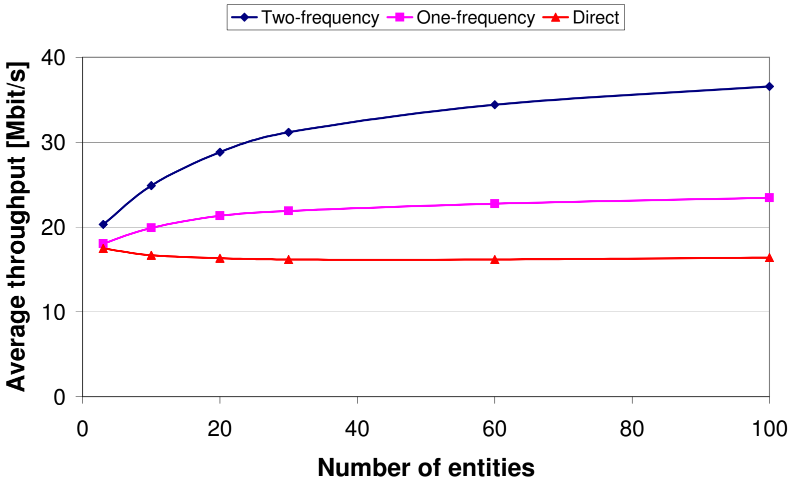


Figure 8.6: Average throughput for single cell as a function of terminals in cell for a uniform traffic fairness scheme, pathloss coefficient $\alpha = 3.2$ and target PER of 1%

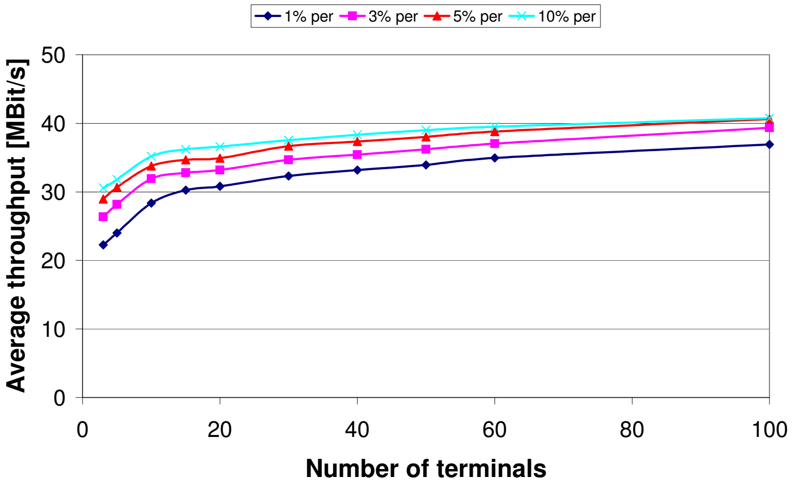


Figure 8.7: Average throughput for single cell as a function of terminals in a cell for a uniform slot size fairness scheme, two-frequency relaying, pathloss coefficient $\alpha = 3.2$, and target PERs of 1%, 3%, 5% and 10%

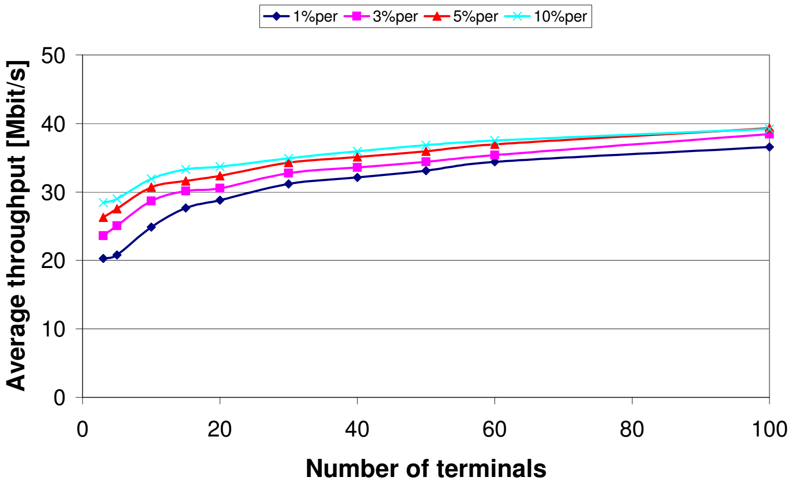
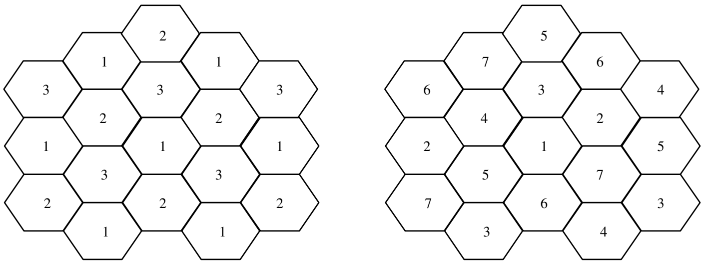


Figure 8.8: Average throughput for single cell as a function of terminals in a cell for a uniform traffic fairness scheme, two-frequency relaying, pathloss coefficient $\alpha = 3.2$, and target PERs of 1%, 3%, 5% and 10%

added as sources of interference to the target cell in consideration and as a source of second frequency for relaying. The relay selection and routing decisions are done similar to the single case, irrespective of the co-channel interference from other cells. But the scheduling, transmission power and rate selections are done in accordance with the interference situation.

Similar to the single cell, a one- and two-frequency relaying modes are utilized using the 3, 7 and 19 frequency reuse patterns. Figure 8.9 and 8.10 show examples of the macro- and micro-frequency assignments for 3 and 7 frequencies. The neighboring cell's frequency is used for the micro-frequency assignment. Sectored-frequency relaying mode is also used to analyze the capacity gain. In this case, the cell is subdivided into six sectors. Each sector uses a recycled frequency from the neighboring cells. An example of such sector relaying is shown in Figure 6.8 for 7-frequency reuse pattern. It should be underlined that these frequencies are used only for relaying purpose close to the APs and the cells keep on using their own macro-frequency plan for their primary communication. In addition, separate relaying frequencies are used to compare the gain in capacity. These separate frequencies are obtained by reassigning the already existing frequencies. For example, for 7-frequency reuse, three frequencies are used for macro-frequency and another three are used for the micro-frequency plan. Similarly for the 19-frequency pattern, seven frequencies are used for each of the macro- and micro-frequency assignments and the results are compared. Note that in all the cases, the *bandwidth per cell* is always constant; different frequency reuse patterns only change the number of different frequency bands in the system at large, but not the allocation per individual cell. Performance differences between the different reuse patterns are thus due to changes in the interference situation.



a. 3- Frequency layout

b. 7- Frequency layout

Figure 8.9: multicell configuration

For the different frequency reuse patters described, the achievable capacity is evaluated by simulation. A total of 61 cells with a radius of 50 m are used in the simulation. The APs are located in a regular hexagonal grid and terminals are uniformly distributed within each cell. The total throughput at the AP is obtained by averaging over the results of 55 different placements. The confidence interval for all results are computed at a 95% confidence level but they are not shown in the figures as they are very small. To protect the results from edge effects, they are computed from the cell at the center.

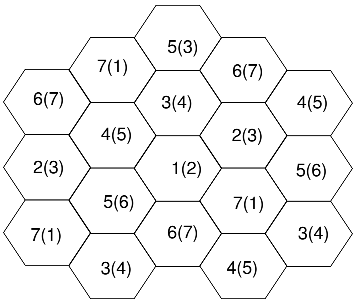


Figure 8.10: An example of macro-micro frequency assignment

Results

A. 7-frequency reuse pattern

Figure 8.11 and 8.12 show the average throughput at the AP for varying pathloss coefficients and target PERs of 1% and 5%. The variation of average throughput obtained by two-frequency relaying for different target PERs and pathloss coefficients is also shown in Figure 8.13. Unlike the single cell results, the average throughput for lower pathloss coefficients is small due to the presence of interference. Since the frequency to be reused is limited to 7, the distance between two co-channel cells that use the macro-frequency pattern is relatively small, making the interference high. When using two-frequency relaying, the source of interference is both from the macro and micro frequency reuse pattern. However, since the communication distance is reduced by the relaying and since the micro frequency plan is used only close to the AP, the interference is reasonably reduced, increased parallelism outweighs the disadvantages. Hence there is up to 19% improvement in average throughput at the AP as shown in Figure 8.14. It is also possible to achieve an improvement of 42% when demand oriented fairness scheme is used to schedule the communication as shown in Figure 8.15. However, the aggregate throughput achieved is less than that of the fairness schemes.

B. 3-frequency reuse pattern

A similar situation is observed for a macro reuse pattern with 3 frequencies as shown in Figure 8.16, 8.17, 8.18 and 8.19. As the frequency reuse distance is small, every transmission is vulnerable to severe co-channel interference. Though the micro frequency reuse scheme amends the situation to some extent, the overall throughput at the AP is considerably less than with 7 frequencies at hand. Nonetheless, when the available reuse frequencies are very few, relaying in two frequencies inevitably gives a better solution. Relaying in two frequency also gives a remarkable result when the scheduling is based on demand oriented fairness scheme. As shown in Figure 8.20, it is possible to achieve about 60% gain in throughput. However, the gain achieved by direct communication only is marked less than the other fairness schemes.

C. 19-frequency reuse

In the case of a 19 frequency reuse pattern, the interfering distance from co-channel cells is relatively large. As shown in Figure 8.21 the one-frequency relaying outperforms the direct communi-

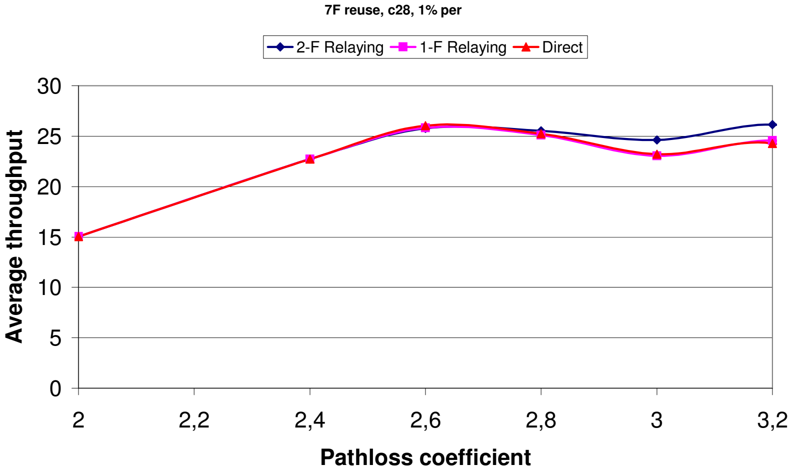


Figure 8.11: Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 1% PER and 7 frequency reuse pattern and uniform slot-size fairness

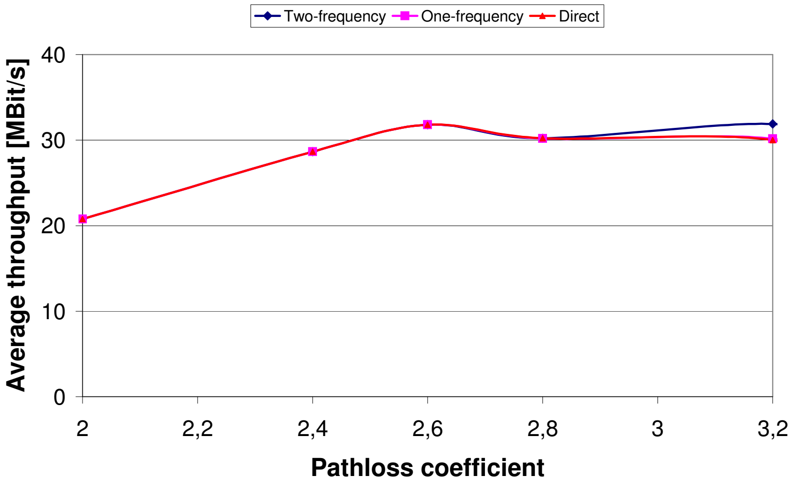


Figure 8.12: Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 5% PER and 7 frequency reuse pattern and uniform slot-size fairness

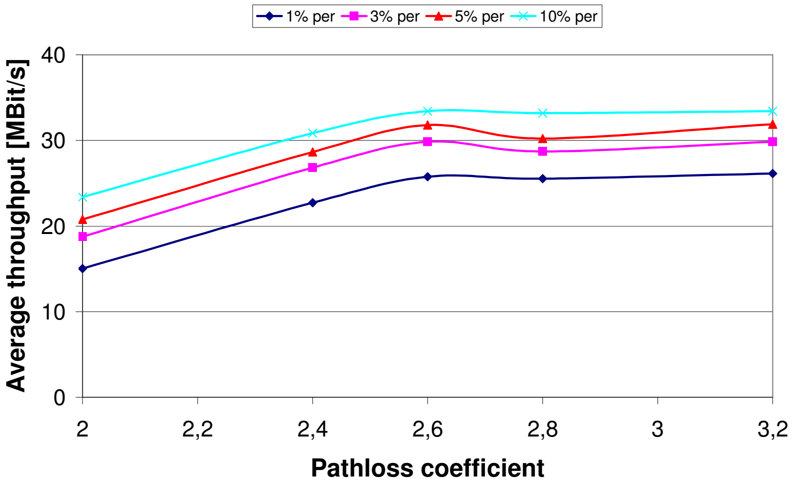


Figure 8.13: Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 2-frequency relaying at 1%, 3%, 5% and 10% PERs and 7 frequency reuse pattern and uniform slot-size fairness

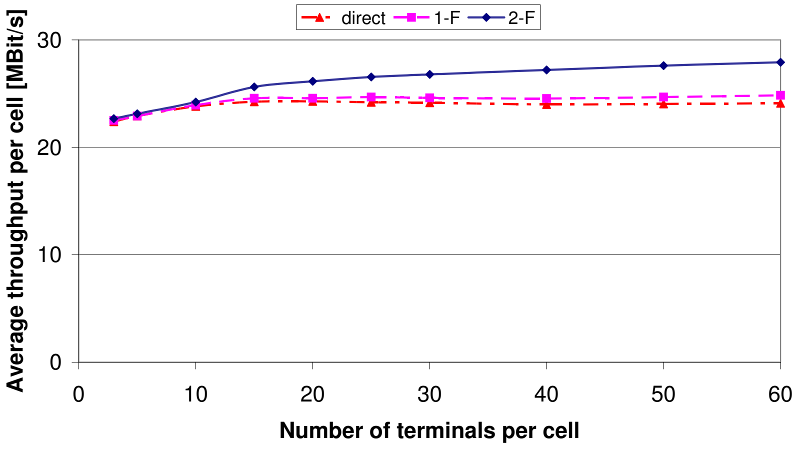


Figure 8.14: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7 frequency reuse pattern and uniform slot-size fairness

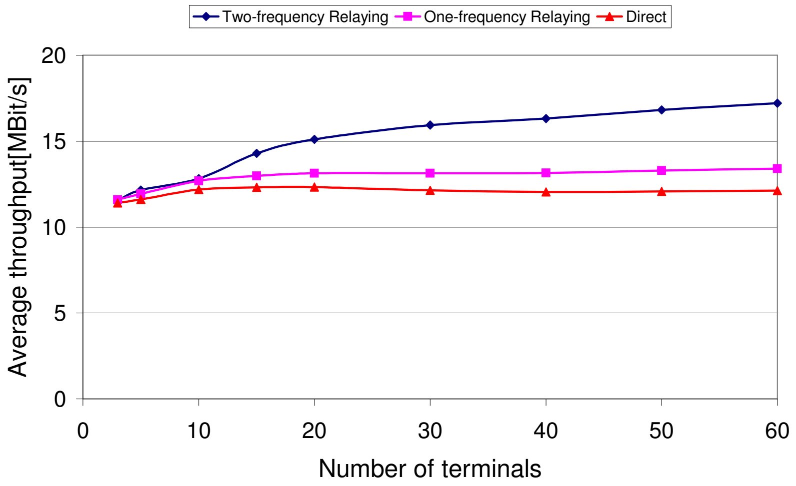


Figure 8.15: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7 frequency reuse pattern and *demand oriented* fairness

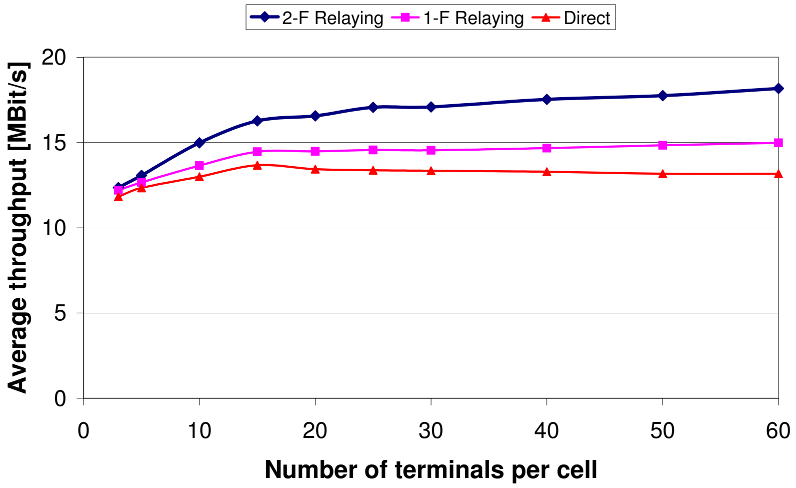


Figure 8.16: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 3 frequency reuse pattern and uniform slot-size fairness

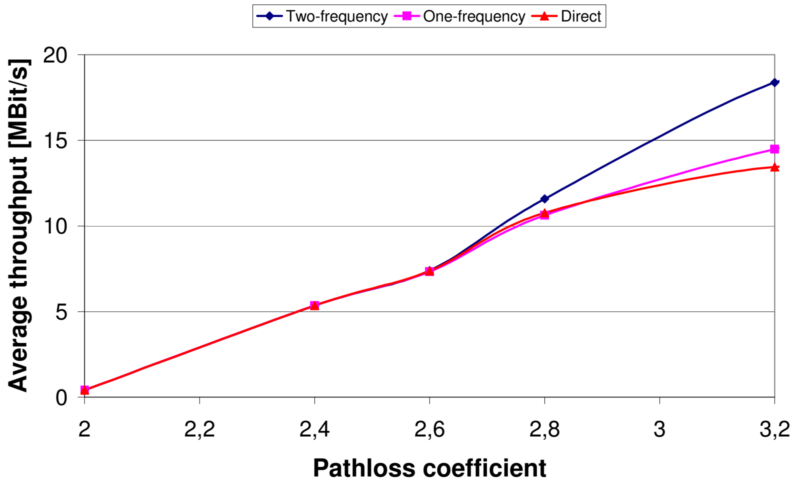


Figure 8.17: Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 1% PER and 3 frequency reuse pattern and uniform slot-size fairness

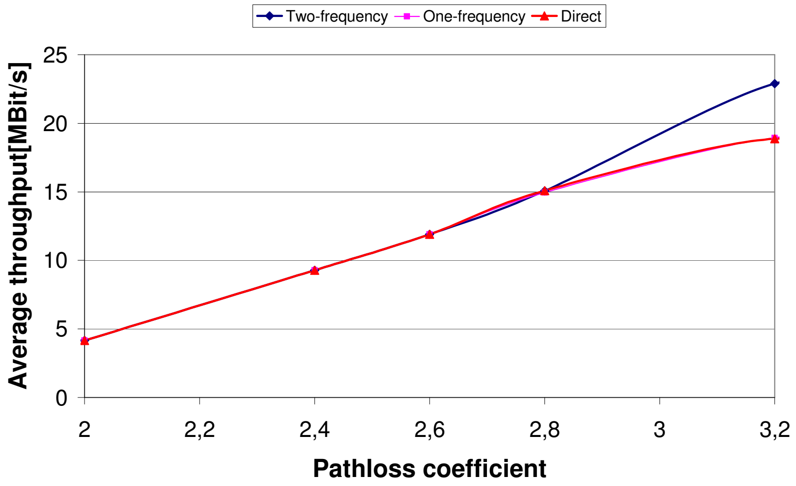


Figure 8.18: Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 5% PER and 3 frequency reuse pattern and uniform slot-size fairness

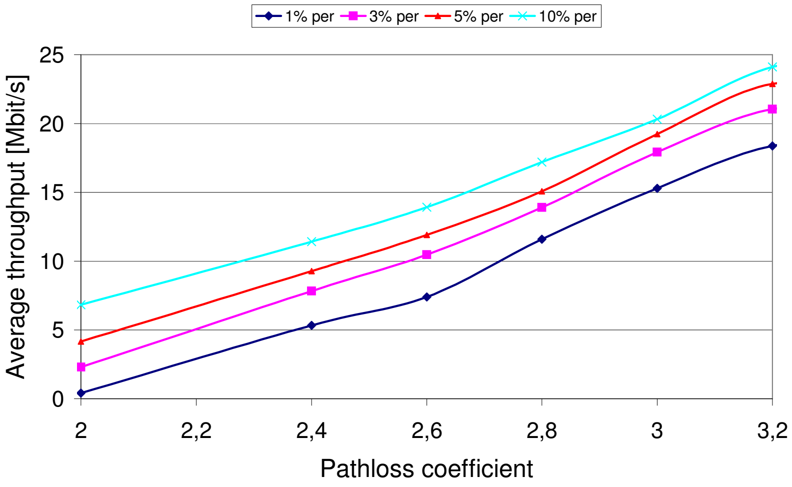


Figure 8.19: Average throughput as a function of $\alpha = 3.2$ for 20 terminals per cell, for 2-frequency relaying at 1%, 3%, 5% and 10% PERs and 3 and uniform slot-size fairness frequency reuse pattern

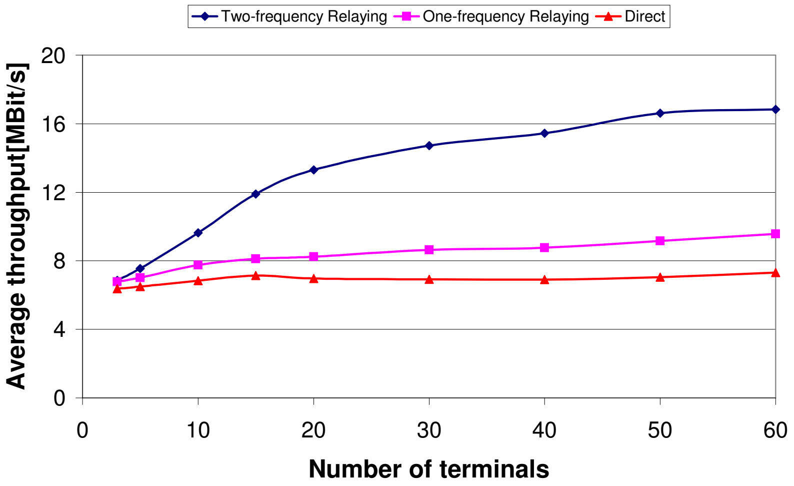


Figure 8.20: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 3 frequency reuse pattern and *demand oriented* fairness

cation. Up to 24% gain is also obtained when a micro-frequency plan with frequency recycling from neighboring cells is applied.

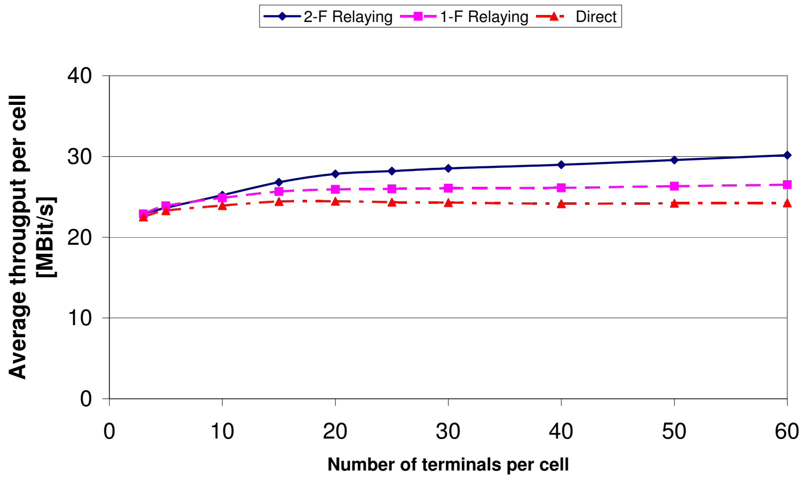


Figure 8.21: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern and uniform slot-size fairness

The variation of the path loss coefficient α and the maximum allowed target packet error rate is shown in Figure 8.22, 8.23 and 8.24. Not surprisingly, relaying is not beneficial for small values of α, as all terminals always can communicate directly with the AP; the initial rising slope is due to better “shielding” of inter-cell interference. Relaying does result in performance gains if α exceeds about 2.8. This is accounted to the packet error rate characteristics for the different pathloss coefficients. As shown in Figure 8.25 the packet error rate becomes significant for most of the modulation rates when α is between 2.6 and 2.8. Consequently, the benefit of relaying is apparent when the packet error rate becomes significant.

Figure 8.24 shows the variation of average throughput of two-frequency relaying with respect to pathloss coefficient α and 1%, 3%, 5% and 10% maximum target packet error rates. Clearly a higher throughput can be attained when the target packet error rate becomes more tolerant but it remains a trade off between a higher throughput and increased number of retransmissions.

Furthermore, Figure 8.26 shows the total average throughput at the AP when uniform traffic size fairness is considered. Although the overall gain achieved by this fairness scheme is lower than that of uniform slot size, it is possible to achieve upto 35% capacity gain using two-frequency relaying and upto 20% one-frequency relaying. Figure 8.27 and 8.28 show the variation of the throughput for different pathloss coefficients and packet error rates. The gain that can be achieved by uniform slot size and demand oriented fairness schemes is almost comparable as shown in Figure 8.29 due to the partial similarity of the scheduling mechanism.

D. Separate relaying frequencies

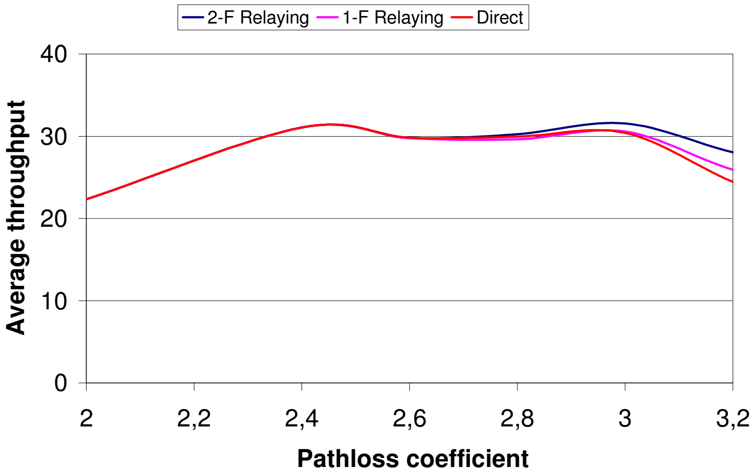


Figure 8.22: Average throughput as a function of varying α, 20 number of terminals, 1% PER, 19 frequency reuse pattern and uniform slot-size fairness

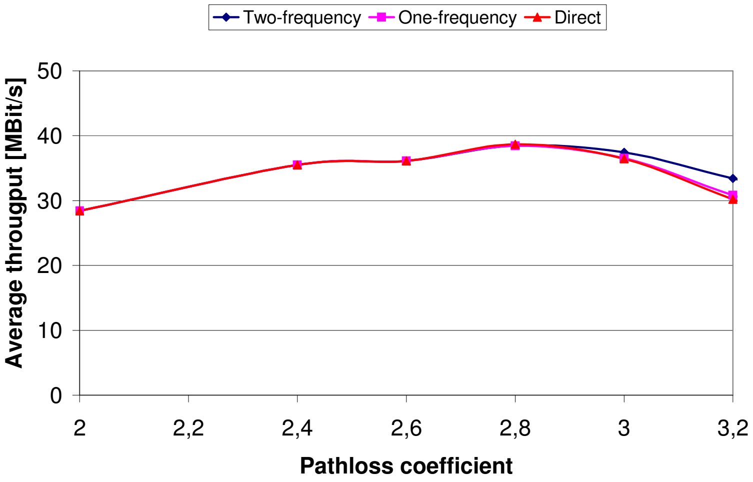


Figure 8.23: Average throughput as a function of varying α, 20 number of terminals, 5% PER, 19 frequency reuse pattern and uniform slot-size fairness

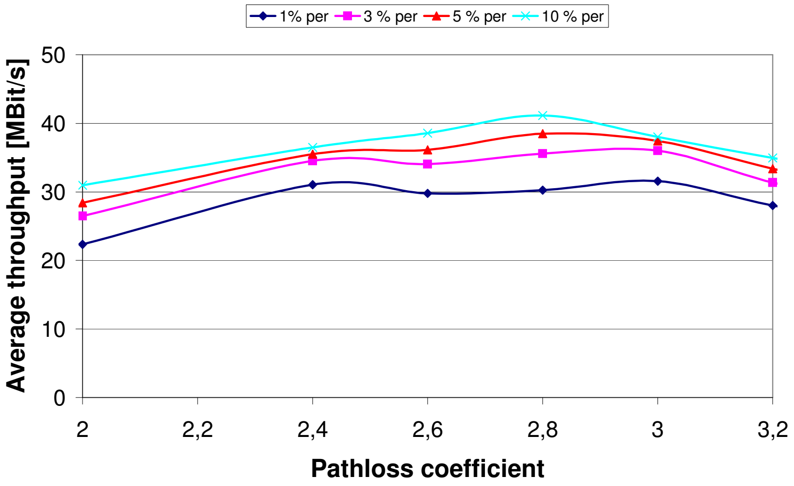


Figure 8.24: Average throughput of 2-frequency relaying as a function of varying α, maximum allowed target PER, 20 number of terminals, 19 frequency reuse pattern and uniform slot-size fairness

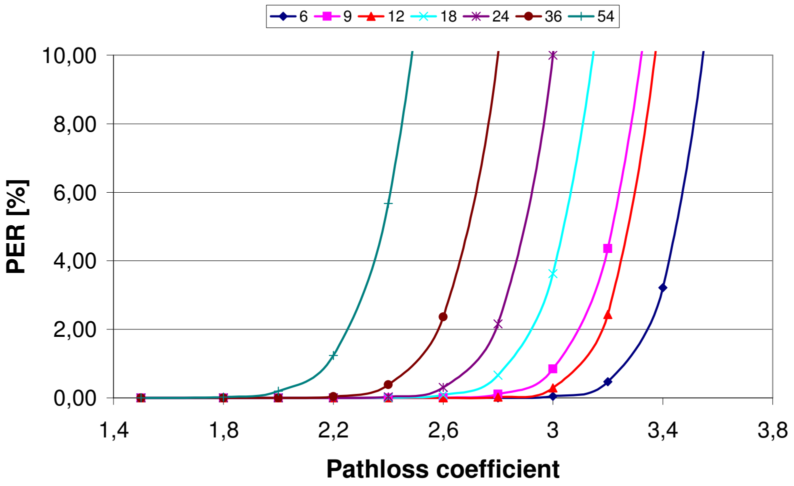


Figure 8.25: Packet error rate variation for the 7 modulation rates used in HiperLAN/2 at a distance of 50 m

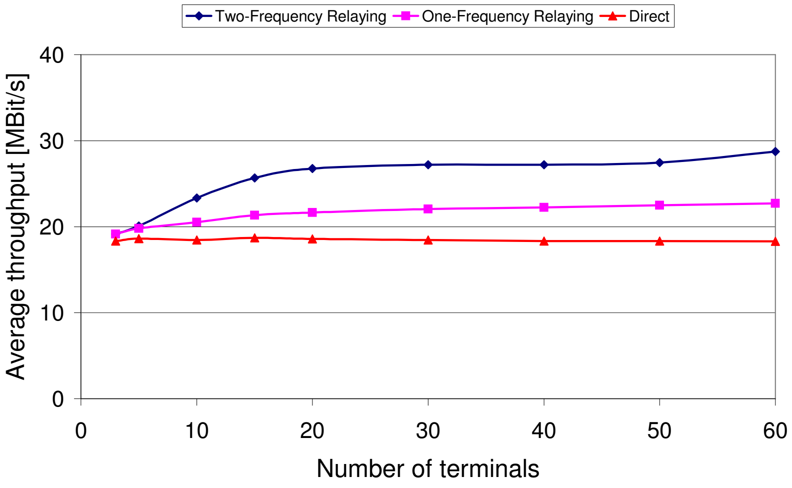


Figure 8.26: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern and *uniform traffic-size* fairness

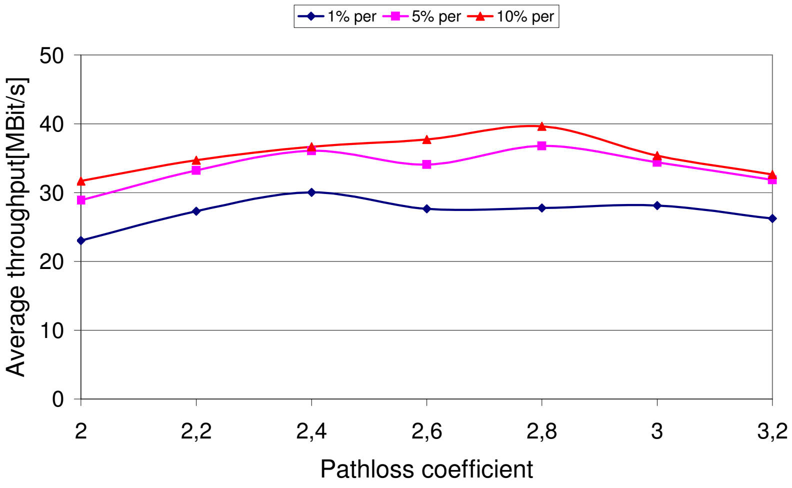


Figure 8.27: Average throughput as a function of varying α for two-frequency relaying mode, 20 number of terminals, 19 frequency reuse pattern, *uniform traffic-size* fairness and for 1%, 5% and 10% PERs

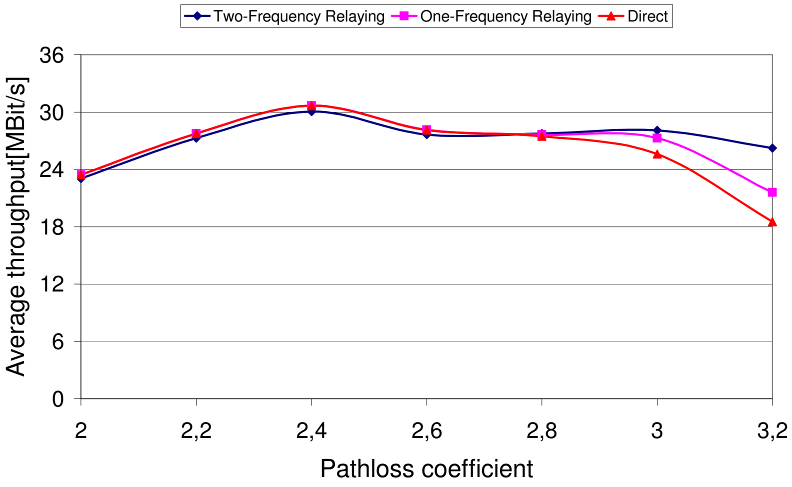


Figure 8.28: Average throughput as a function of varying α, 20 number of terminals, 1% PER, 19 frequency reuse pattern and *uniform traffic-size* fairness

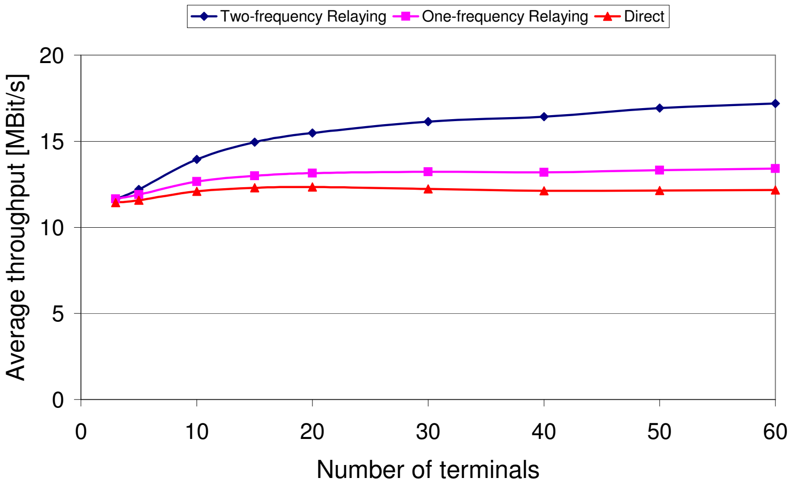


Figure 8.29: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern and *demand oriented* fairness

Comparing the results in Figures 8.14 and 8.21 for the 7- and 19- frequency patterns, the again achieved by adding 12 frequencies is actually small. Hence, these frequencies might perhaps be better used to support relaying. The interesting situation is when a 7 frequency pattern is used for the macro frequency reuse pattern and additional 7 frequencies (i.e. 14 frequencies in the system as a whole) are available for the micro reuse pattern as described in Section 6.6. As the recycled frequencies are not used from the immediate neighboring cells, the reuse distances are increased, which evidently brings about capacity improvement by reducing the level of interference as shown in Figure 8.30. Also, Figure 8.31 compares this situation with that of the 19 frequency reuse pattern. Up to 40 % improvement can be obtained by the two-frequency relaying when the new micro frequency reuse pattern is used and yet 5 frequencies can be saved compared to the standard 19 frequency reuse pattern. This represents a good incentive to have a smaller frequency plan with a better figure of merit. Likewise, for uniform traffic size fairness scheme up to 40% performance improvement is achieved using separate relaying frequencies as shown in Figure 8.32. The throughput variation for different pathloss coefficients and packet error rates is also shown in Figures 8.27 and 8.28. It is to be noted that the gain achieved in all the cases is at no cost, without requiring any additional hardware for the existing technology.

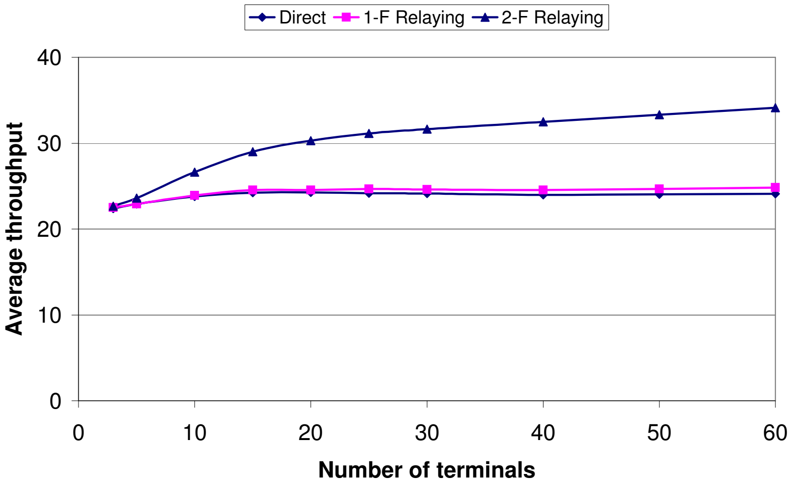


Figure 8.30: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7+7 frequency reuse pattern and uniform slot-size fairness

Similarly, when 3 frequency is used for macro-frequency reuse plan and another 3 frequency for the micro-frequency, the interference condition is harmonized and relatively higher throughput at the AP is obtained compared to the 3 frequency reuse plan. However, the achievable gain is less than the 7 frequency reuse plan without separate relay frequency. Figure 8.35 and 8.36 depict the situation.

E. Sectored frequencies relaying

Multi-frequency relaying is used within the cell by dividing the cell into six sectors and assigning different recycled frequencies from neighboring cells. The result of this sectored frequency relaying

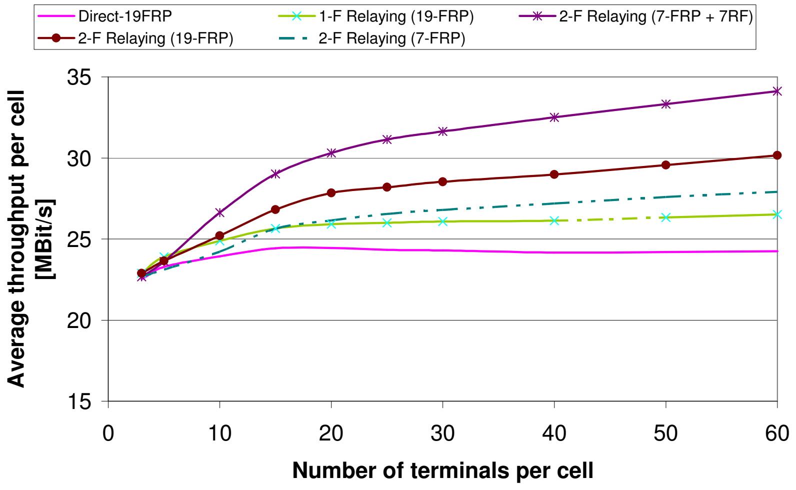


Figure 8.31: Comparisons of average throughput as a function of terminals per cell for different frequency reuse patterns, 1% PER and $\alpha = 3.2$ for uniform slot-size fairness

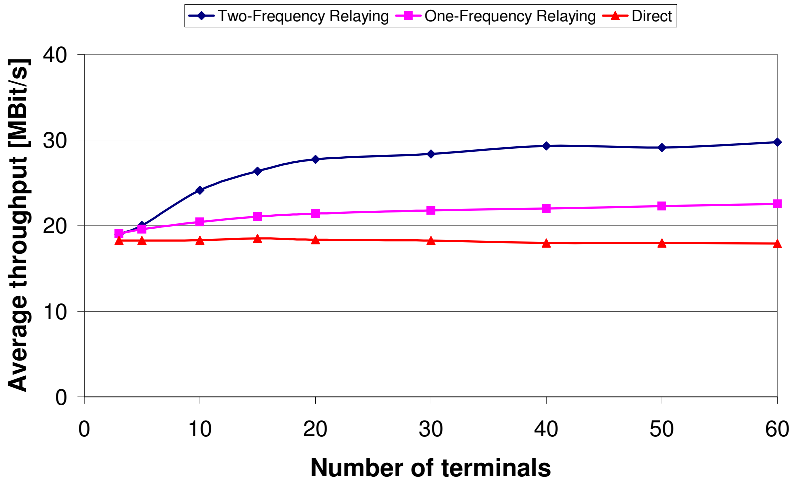


Figure 8.32: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7+7 frequency reuse pattern and *uniform traffic-size* fairness

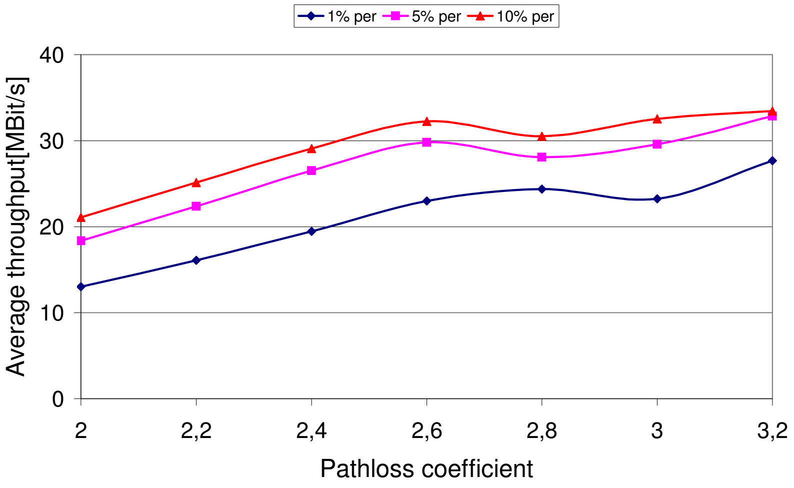


Figure 8.33: Average throughput as a function of varying α for two-frequency relaying mode, 20 number of terminals, 7+7 frequency reuse pattern, *uniform traffic-size* fairness and for 1%, 5% and 10% PERs

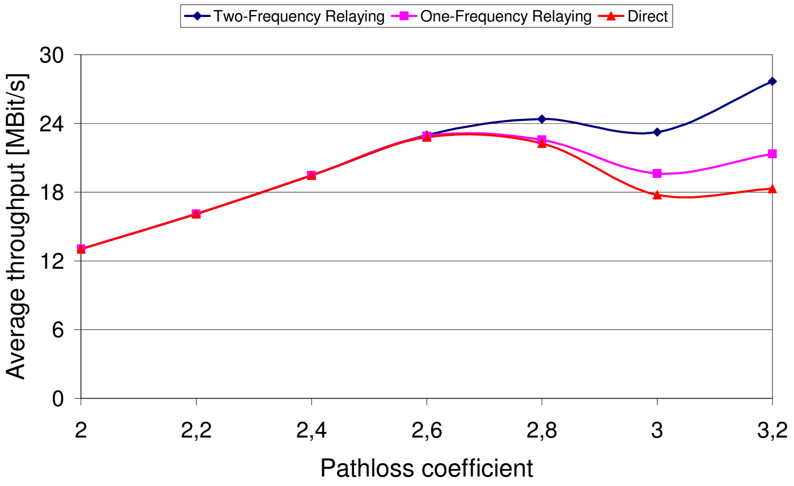


Figure 8.34: Average throughput as a function of varying α, 20 number of terminals, 1% PER, 7+7 frequency reuse pattern and *uniform traffic-size* fairness

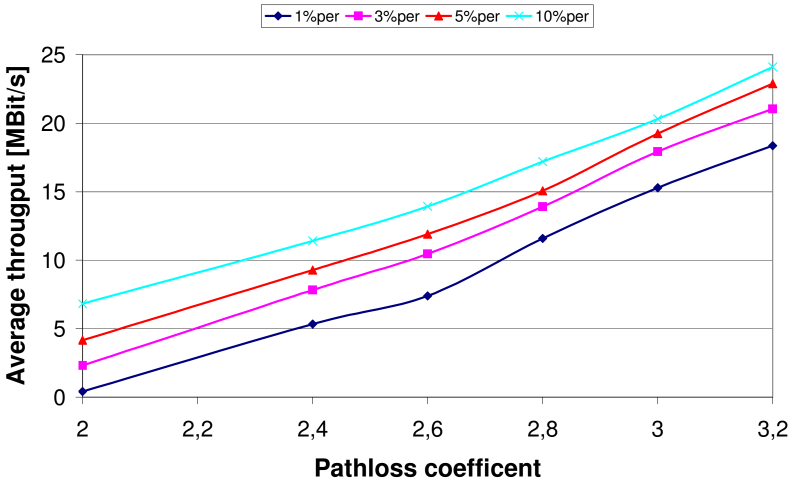


Figure 8.35: Average throughput as a function of varying α, 20 number of terminals for two-frequency relaying mode, 3+3 frequency reuse pattern and uniform slot-size fairness

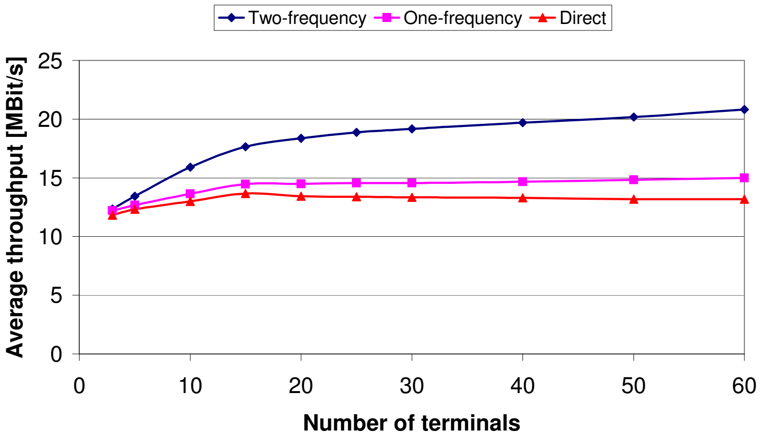


Figure 8.36: Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$, 3+3 frequency reuse pattern and uniform slot-size fairness

is shown in Figure 8.37, 8.38 and 8.39. In 19 frequency reuse pattern case, the reuse distance for the relaying frequencies obtained from the neighboring cells is relatively large. The ensuing interference from sectored frequency relaying is not that critical to reduce the gain and hence comparably high throughput is obtained with that of two-frequency relaying. For sector relaying in 3 and 7 frequencies, the gain achieved is less than that of direct communication due to the excessive interference arising from multi-frequency transmissions. If separate frequencies are used for relaying, the result shows improvement. Nevertheless, the overall gain is not significant if the multi-frequency switching overhead is also taken into consideration.

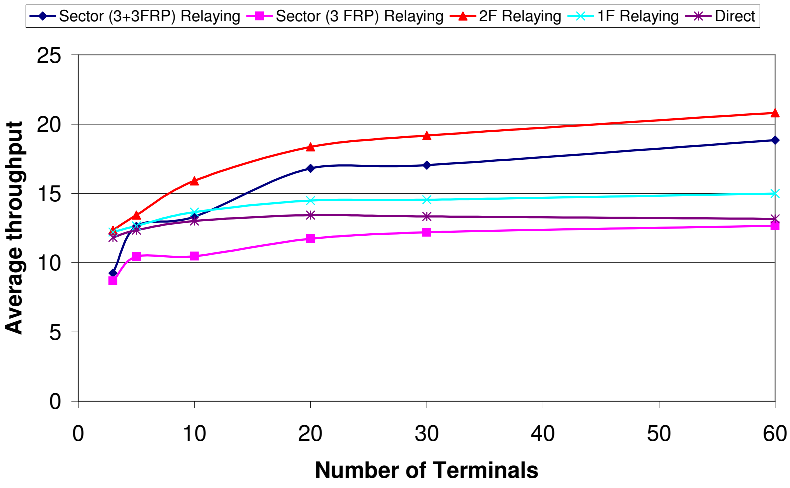


Figure 8.37: Average throughput of sectored-frequency relaying as a function of varying terminals, maximum allowed target PER, $\alpha = 3.2$ for 3 frequency reuse pattern and uniform slot-size fairness

Summary of the Results

In this chapter the simulation results for the proposed rate-adaptive, two-frequency relaying and fair scheduling algorithms are presented. The results proved the potential benefit of relaying to improve the capacity of the cellular network considered. The level of improvement depends on:

- The frequency reuse patterns: 3-, 7- and 19-frequency. In addition, when separate frequencies are used for relaying, i.e., 3 frequencies for macro and additional three for micro frequency (3+3) and 7 for macro and additional 7 for micro (7+7) frequency patterns.
- The relaying mode used: direct, one-frequency relaying, two-frequency and sector-based frequency relaying modes.
- The fairness scheme chosen to schedule the TDMA communication frame: uniform slot size, uniform traffic size and demand oriented schemes.

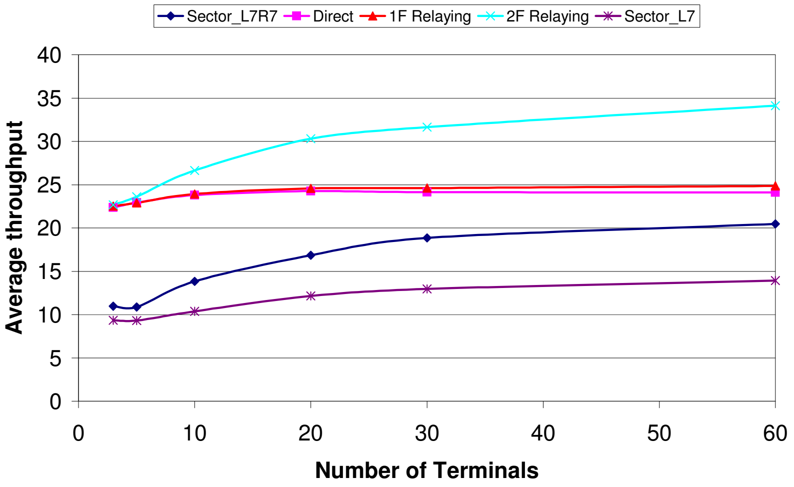


Figure 8.38: Average throughput of sectored-frequency relaying as a function of varying terminals, maximum allowed target PER, $\alpha = 3.2$ for 7 frequency reuse pattern and uniform slot-size fairness

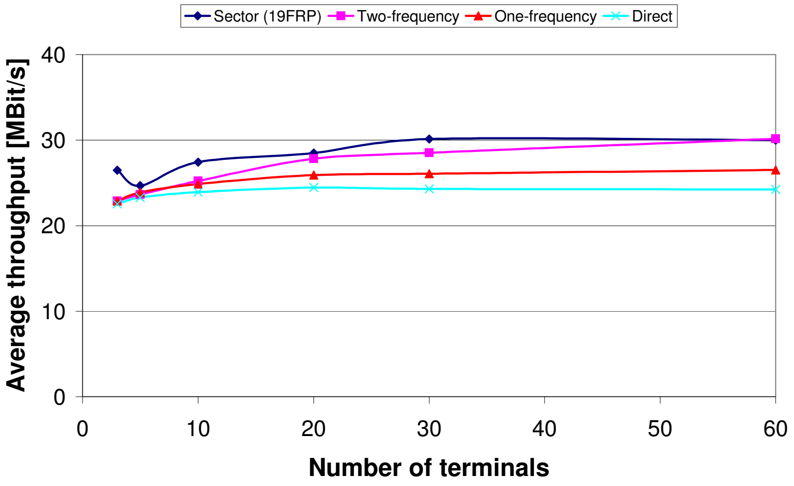


Figure 8.39: Average throughput of sectored-frequency relaying as a function of varying terminals, maximum allowed target PER, $\alpha = 3.2$ for 19 frequency reuse pattern and uniform slot-size fairness

Comparing the analytical results achieved in Chapter 5 and the simulation results presented here, they both substantiate the benefit of relaying. And it is possible to see that the results from the simplified models are scaled to larger scenarios. But the gain achieved in the simulation result is not as in the analytical cases since there are various limiting factors in practical considerations. Nonetheless, relaying is always superior to the direct communication and brings significant capacity improvement. For the multi-hop cellular network considered, it is shown that up to 60% capacity gain can be achieved. In both customer oriented and service provider oriented analyses, the two-frequency relaying provides the best performance for the cellular wireless network. Among the fairness schemes, the uniform slot size fairness renders the best performance, as well.

As indicated in the results, the benefit of relaying is more significant when the pathloss coefficient is larger than 2.8 and when the number of users is relatively large. The result confirms to the anticipated relaying application area– such as in exhibition halls, airport and supermarkets where the pathloss coefficient is typically between 3 and 4 [84], and where many users communicate with each other using wireless networks.

It is to be recalled that the key result of GUPTA and KUMAR [41] for the capacity of wireless networks is that for n randomly located nodes in a unit disk with a uniform traffic pattern, the maximum achievable per node throughput is $O(\frac{1}{\sqrt{n}})$ and as the number of nodes n per unit area increases, the per node throughput goes to zero. This result the best performance achievable even allowing for optimal scheduling, routing and relaying of packets in the network. However, with practical considerations, our results show an overall increase in system capacity. Interestingly, the per-node throughput behaves approximately like $\frac{1}{\log(n)}$. And for large number of nodes, the per node throughput will **not** actually go to zero but becomes a constant. An example of this situation is shown in Figure 8.40 for different frequency reuse patterns.

Moreover, the algorithms are practical as they are not computationally intensive, they can be implemented as iterative online algorithms and are based on information that can be provided by real systems with acceptable overhead. As no additional infrastructure is necessary and also the requirements on the individual terminals are quite modest, this relaying approach can provide a simple and cheap solution to add capacity to a wireless system, particularly in highly loaded networks.

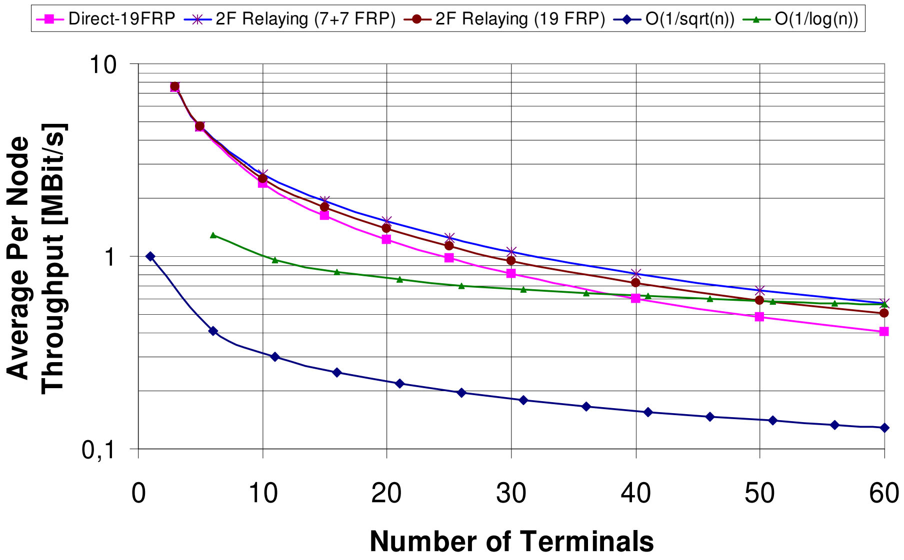


Figure 8.40: Per-node average throughput as a function of terminals per cell for different frequency reuse patterns and 1% PER, $\alpha = 3.2$

Chapter 9

Impact of Mobility

9.1 Introduction

In the discussion of wireless ad hoc networks in Chapter 4, we have seen the fundamental study of GUPTA and KUMAR [41] on the capacity of fixed ad hoc wireless networks. Their result showed that as the number of nodes per unit area n increases, the per-node throughput decreases approximately like $\frac{1}{\sqrt{n}}$. This was the best performance achievable even allowing for optimal scheduling, routing and relaying of packets in the network. The result puts the feasibility of large ad hoc networks in question as the per node throughput actually goes to zero for large number of nodes.

GROSSGLAUSER and TSE [40] introduced mobility into the model that GUPTA and KUMAR [41] used and considered the situations where nodes move independently around the network. Unlike in [41], their main result showed that the average long term per node throughput can be kept constant even as the number of nodes per unit area increases. In other words, there is no significant loss in per node throughput in terms of growth rate as a function of the number of nodes. A similar result was also achieved when nodes have limited mobility pattern and are constrained to move only in a one-dimensional space [20].

These performance improvements are attained through the exploitation of the time variation of the nodes' channel due to mobility. They attempted to split the packets of each source node to as many different nodes as possible. These nodes then serve as mobile relays and whenever they come close to the final destination, they hand the packets off to the final destination. The basic idea behind is that since there are many different relay nodes, the probability that at least one is close to the destination is significant and hence a better capacity can be achieved. However, the caveat of the result is that the attained long term throughput is averaged over the time scale of node mobility and hence delay of large order will be incurred.

The issue of whether mobility can be used to obtain a close-to-optimal throughput as in [40] which at the same time guarantees low delay is addressed by BANSAL and LIU [8]. They considered a mobile ad hoc network with n static nodes uniformly distributed in a disk, and m mobile nodes who move according to some defined mobility model. The sender-receiver pairs are chosen among the static nodes and the m mobile nodes assist in handing off packets to the destination. Their main result shows that there exist a constant $c > 0$ for a maximum available bandwidth W, such that the per node throughput is $c\frac{W \min(m,n)}{n \log^3 n}$ and the maximum delay incurred by the packet is at most $\frac{2d}{v}$, where d is the diameter of the network and v is the velocity of the mobile nodes.

In summary, due to the long-range direct communication between many sender-receiver pairs in fixed wireless networks, the subsequent interference in the system primarily limits the capacity of the network. One possible way to surmount this fundamental limitation is multi-hop communication. And the studies in [8, 20, 40] showed the possibility of further improving the capacity of wireless network by exploiting the mobility of nodes. These results are for purely ad hoc wireless networks and it is still an open issue if their assertion is also true to multi-hop cellular networks.

In this chapter two issues are addressed regarding mobility: First, the hypothesis that mobility can improve the capacity of a wireless network is examined in a multi-hop cellular network. This is done by analyzing the proportion of terminals in a cell that can take advantage of mobility to get a higher throughput. Secondly, the novel approaches used in this work to improve the capacity of cellular networks are under the premise that the terminals are stationary. But one of the charms of wireless communication is the ability of the users to move around while communicating. When there is mobility in the network, obviously the topology changes frequently and the routing should be updated to maintain the throughput achieved. The second part thus discourses issue related to the topology change due to mobility.

The remainder of the chapter is organized as follows: The mobility model to be used is described in Section 9.2. The relay utilization of a cellular network in the presence of mobility is analytically presented in Section 9.3. The simulation set up to study the effect of topology change in the capacity of mobile cellular network is presented in Section 9.4. Finally the corresponding results and discussion are presented in Section 9.5 and 9.6.

9.2 Mobility Model

The movement pattern of nodes plays an important role in performance analysis of mobile and wireless networks. Once the nodes are initially placed, the mobility model dictates how the nodes move within the network. A variety of mobility models have been presented for wireless networks [10, 15, 51, 58]. These models vary widely in their movement characteristics.

Among the various models, the popular Random Walk model [15] is used in this thesis to assess the impact of mobility on the static routing and scheduling algorithm. In this mobility model, terminals move from their current location to a new location by randomly choosing a direction and speed to travel. The new speed and direction are both chosen from pre-defined ranges $[speedmin, speedmax]$

and $[0, 2\pi]$ respectively. Once terminals choose new speed and direction, they move for a specified time t. The current speed and direction is independent of its past speed and direction.

In this experiment, terminals can move within the cell they belong but they are not allowed to cross a boundary and join a new cell. Thus, handover is not an issue here. Instead, if a terminal reaches at a boundary of a cell, it bounces off with an angle determined by the incoming direction. The terminal then continues to move along this new path. Figure 9.1 shows the traveling pattern of a mobile terminal within a cell of radius 50m whose AP is located at the origin. The terminal chooses a speed between 0 and 10m/s and a direction between 0 and 2π. It is allowed to move at the chosen speed and direction for 10 steps.

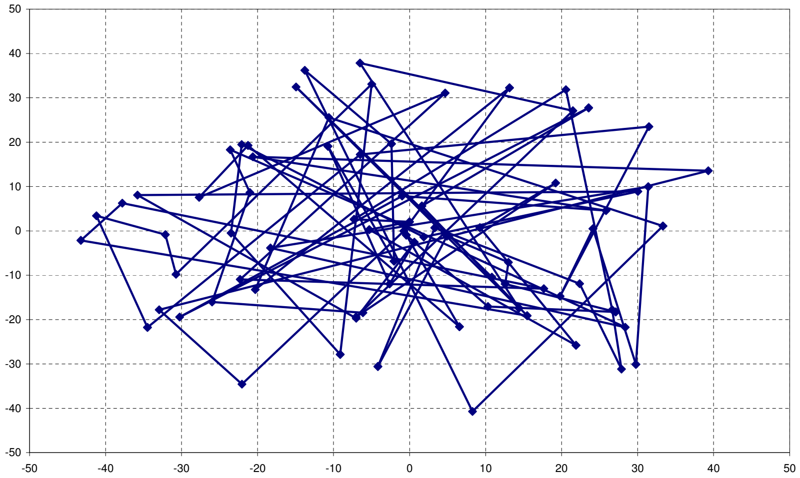


Figure 9.1: Traveling pattern of a mobile terminal with a random walk mobility model

To evaluate the effect of the rate of mobility (speed of the mobile terminal) on the system performance, a modified Random Walk Mobility model is also used. In this case, every terminal is assigned the same speed but chooses direction randomly from the range $[0, 2\pi]$ at the beginning. Then the assigned speed value is increased in steps in the range $[speedmin, speedmax]$.

9.3 Analysis of Relay Utilization due to Mobility

The rationale for relaying is providing shorter communication distance which allows sender terminals to select higher transmission rates to send their traffic. The hypothesis that mobility increases capacity is in the sense that there will be a mobile terminal which comes closer to the sender terminal at some instant of time and is able to receive the traffic with the highest possible data rate. The mobile terminal then relays the traffic when it arrives in the vicinity of the destination terminal at a later time.

The following two questions give an insight to justify the hypothesis for multi-hop cellular networks as well, and to examine the extent to which mobility contributes in improving the system capacity:

- how often is a mobile terminal available which comes closer to the sender terminal and relays the traffic to the destination within some delay threshold time?
- how dense should the network be in order to have a significant improvement?

This section attempts to answer the above questions by studying a single-relay-hop cellular network with terminals capable of moving within the cell.

9.3.1 Network Model

Consider the cellular network model shown in Figure 9.2. The network consists of an AP, a far terminal S and a mobile terminal MT. The terminals are assumed to be in the radio coverage of the cell. R is the radius of the cell and r is the radius of higher transmission rate range. A uniform omni-directional radio propagation is assumed, and the derivations shown here are strictly geometric.

Let us assume that the far terminal S has traffic to send to the AP but it can only send with the maximum transmission power and slowest data rate. Otherwise terminal S needs to use an intermediate terminal to send its traffic with higher transmission rate. The aim now is to determine the probability of finding a mobile relay terminal within the higher transmission range of terminal S at time t_1 such that this relay terminal moves in the direction of the AP and reaches within the higher transmission range of the AP at time t_2 for $t_2 - t_1 \leq T_D$, where T_D is some delay threshold time.

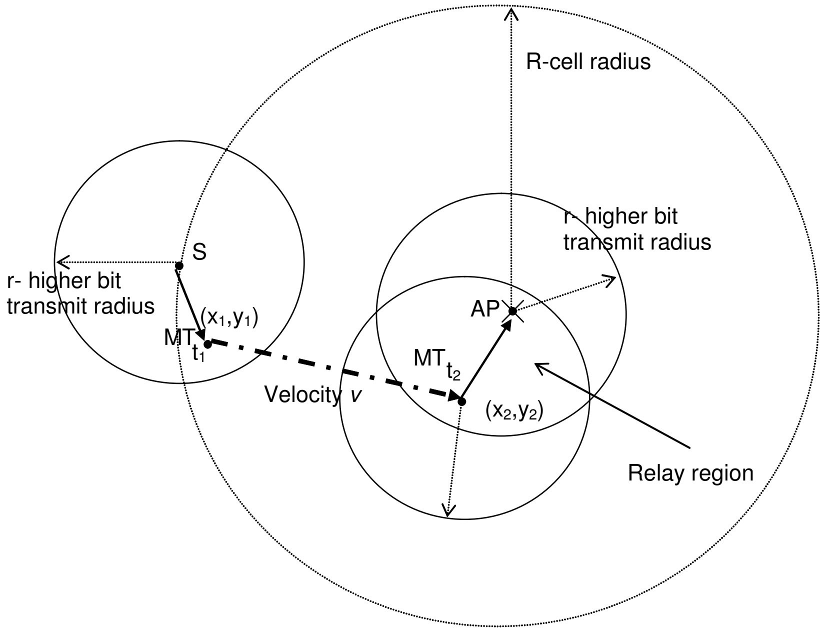


Figure 9.2: Simple single-relay-hop network with relay region

9.3.2 Probability Derivation for Finding a Mobile Relay Terminal

Let the AP is located at $(0,0)$ and the source terminal S at (x_s, y_s). A single-relay-hop with a higher transmission rate is possible if the following conditions are met:

1. the distance between S and MT at time t_1 is less than r
2. the time needed for MT to reach within the transmission radius r of AP (in the relay region) from the $S's$ transmission radius range should not exceed the delay treshold, i.e., $t_2 - t_1 \leq T_D$
3. the distance between the AP and MT at time t_2 is less than r.

Assume that the mobile terminal MT moves within the cell with Random Walk mobility model discussed in Section 9.2. To see the extent to which the mobility further improves the system capacity, we need to determine the probability of finding such terminals in the network. Following this, the aim of this section is to:

1. determine the probability of finding a mobile terminal within the higher transmission range of terminal S at time t_1 and this mobile terminal relays the traffic to the AP when it comes closer to the higher transmission range of the AP at time t_2 such that $t_2 - t_1 \leq T_d$, where T_d is some threshold delay time.
2. determine the expected proportion of terminals that satisfy the condition in 1.

Derivation

Assume that the transmission range of terminal S be in A(S,r), area with radius r and center at S and let A(O,R) be the coverage area of the cell centered at O with radius R. Let there exist N mobile terminals uniformly distributed in the cell and independent of each other. Let ζ be the event that there is a mobile terminal in A(S,r). At time t=0, the probability that at least one of the N-1 terminals is with in A(S,r) is given by:

$$\Pr(\zeta) = 1 - \left(1 - \frac{\lambda(A(S,r))}{\lambda(A(O,R))}\right)^{N-1} \tag{9.1}$$

where $\lambda(\cdot)$ is the surface of the transmission coverage. If the event ζ is not satisfied, the sender terminal A either sends its traffic with the slowest data rate or waits for some time until the event ζ is true.

Now let us assume that ζ is true and choose a single random mobile terminal at position Y_o from A(S,r) in time t=0. At any time t, the mobile terminal moves at a randomly chosen direction and speed v as:

$$Y_t = Y_0 + vt \cdot A_\alpha e_1 \tag{9.2}$$

where Y_0 is uniformly distributed in A(S,r), e_1 is a direction vector and A_α is a random angle with uniform distribution given by

$$A_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \tag{9.3}$$

According to the random direction shown, the mobile terminal x can either reach the higher transmission range of the AP for a finite time τ or it will never come close to the AP transmission range in one step as shown in case 1 and 2 of Figure 9.3.

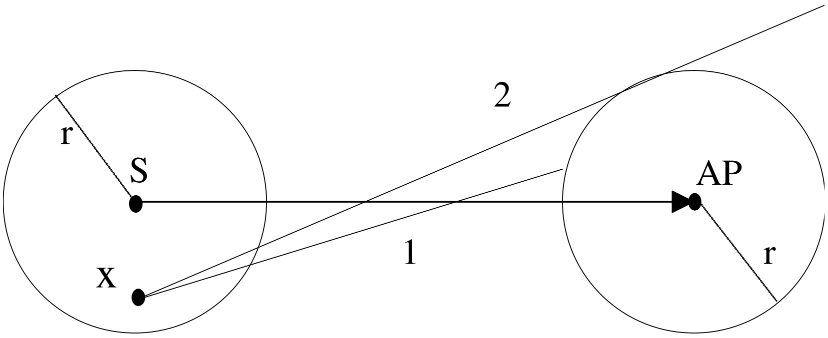


Figure 9.3: Possible directions that a mobile terminal x can move towards the AP

In other words:

$$\tau = \inf\left\{t : Y_t \in A(AP, r)\right\} \tag{9.4}$$

Thus the probability that a mobile terminal is within A(S,r) moves at a speed v and reaches in A(O,r), higher transmit radius r range of the AP, within a finite time is:

$$\Pr\left[\tau < t\right] = \frac{1}{A(S,r)} \int\limits_{A(S,r)} \Pr(\tau < t | Y_0 = x) dx \tag{9.5}$$

which is given as:

$$\Pr\left[\tau < t\right] = \begin{cases} 0 & vt \le |x-O| - r \\ \frac{1}{\pi}\arccos\left(\frac{(vt)^2 + |x-O|^2 - r^2}{2vt|x-O|}\right) & |x-O| - r \le vt < \sqrt{|x-O|^2 - r^2} \\ \frac{1}{\pi}\arcsin\left(\frac{r}{|x-O|}\right) & vt \ge \sqrt{|x-O|^2 - r^2} \end{cases} \tag{9.6}$$

Let the AP is at located (0,0) and the sender terminal S is at (s,0) as in the coordinate shown in Figure 9.4.
The probability is then:

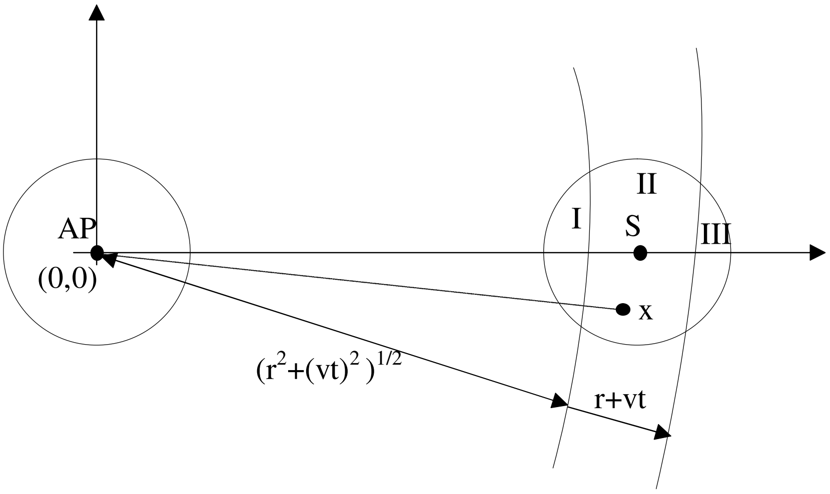


Figure 9.4: The mobile terminals in rectangular coordinate

$$\Pr[\tau < t] = \frac{1}{A(S,r)} \int\limits_{A(S,r)} \Pr(\tau < t | Y_0 = x) dx$$

$$= \frac{1}{A(S,r)} \int_{A_I} \frac{1}{\pi} \text{arcsin} \left(\frac{r}{|AP - x|} \right) d\alpha \quad + \quad \frac{1}{A(S,r)} \int_{A_{II}} \frac{1}{\pi} \text{arcsin} \left(\frac{(vt)^2 + |AP - x|^2 - r^2}{2vt\,|AP - x|} \right) d\alpha \tag{9.7}$$

Changing in the coordinate to polar as shown in Figure 9.5:

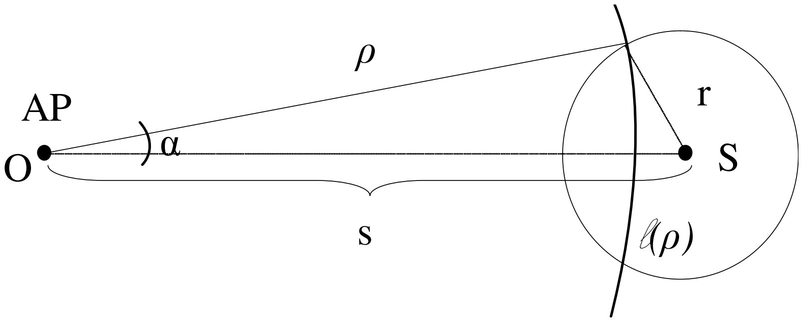


Figure 9.5: The mobile terminals in polar coordinate

$$= \frac{1}{A(S,r)} \int_{A_I} \rho \frac{1}{\pi} \text{arcsin} \left(\frac{r}{\rho} \right) l(\rho) d\rho \quad + \quad \frac{1}{A(S,r)} \int_{A_{II}} \rho \frac{1}{\pi} \text{arcsin} \left(\frac{(vt)^2 + \rho^2 - r^2}{2vt\rho} \right) l(\rho) d\rho \tag{9.8}$$

From Figure 9.5, $l(\rho)$ is given by:

$$\alpha = \arcsin\left(\frac{s^2+\rho^2-r^2}{2s\rho}\right) \tag{9.9}$$

and

$$l(\rho) = 2\rho \arcsin\left(\frac{s^2+\rho^2-r^2}{2s\rho}\right) \tag{9.10}$$

Rewriting the probability expression:

$$\begin{aligned} = \quad & \frac{1}{A(S,r)}\int_{\min\left(s-r,\sqrt{(vt)^2+r^2}\right)}^{\min\left(\sqrt{(vt)^2+r^2},s+r\right)} \rho\frac{1}{\pi}\arcsin\left(\frac{r}{\rho}\right) l(\rho)d\rho \quad + \\ & \frac{1}{A(S,r)}\int_{\min\left(\sqrt{(vt)^2+r^2},s-r\right)}^{\min\left(r+s,vt+r\right)} \rho\frac{1}{\pi}\arcsin\left(\frac{(vt)^2+\rho^2-r^2}{2vt\rho}\right) l(\rho)d\rho \end{aligned} \tag{9.11}$$

Finally the probability that at least one of the N-1 terminals is with in A(S,r) and reaches in the higher transmission range of the AP is:

$$\frac{1}{A(S,r)}\int_{A(S,r)} \Pr(\tau < t|Y_0 = x)dx \cdot \left(1-\left(1-\frac{\lambda(A(S,r))}{\lambda(A(O,R))}\right)^{N-1}\right) \tag{9.12}$$

which is:

$$\begin{aligned} = \quad & \frac{1}{A(S,r)}\int_{\min\left(s-r,\sqrt{(vt)^2+r^2}\right)}^{\min\left(\sqrt{(vt)^2+r^2},s+r\right)} \rho\frac{1}{\pi}\arcsin\left(\frac{r}{\rho}\right) l(\rho)d\rho \quad + \\ & \frac{1}{A(S,r)}\int_{\min\left(\sqrt{(vt)^2+r^2},s-r\right)}^{\min\left(r+s,vt+r\right)} \rho\frac{1}{\pi}\arcsin\left(\frac{(vt)^2+\rho^2-r^2}{2vt\rho}\right) l(\rho)d\rho \cdot \left(1-\left(1-\frac{\lambda(A(S,r))}{\lambda(A(AP,R))}\right)^{N-1}\right) \end{aligned} \tag{9.13}$$

Expected number of terminals

The next step is to determine the expected proportion of mobile terminals which can satisfy the conditions stated earlier and which can bring significant improvement in the system performance by use of relaying. Let $x_1, \ldots, x_i, \ldots, x_{N-1}$ are terminals that are independent and uniformly distributed with in the cell. The expected number of terminals is given by:

$$E\sum_{i=1}^{N-1} P_\zeta(x_i) = \sum_{i=1}^{N-1} E(P_\zeta(x_i)) = N \cdot \frac{\lambda(A(S,r))}{\lambda(A(O,R))} \tag{9.14}$$

where P_ζ is the probability of mobile terminals satisfying the condition and is expressed as:

$$P_\zeta(x_i) = \begin{cases} 1 & if \quad x \in A(S,r) \\ 0 & else \end{cases} \tag{9.15}$$

9.3.3 Numerical Results

In this section the numerical solution of the above probability integral is presented for various values of N, vt, the distance covered in time t by the randomly chosen velocity v, and D, the ratio of cell radius to higher transmission rate radius. The results are on the assumption that there is only single-relay-hop transmission and vt is the distance covered by the mobile relay without changing its direction.

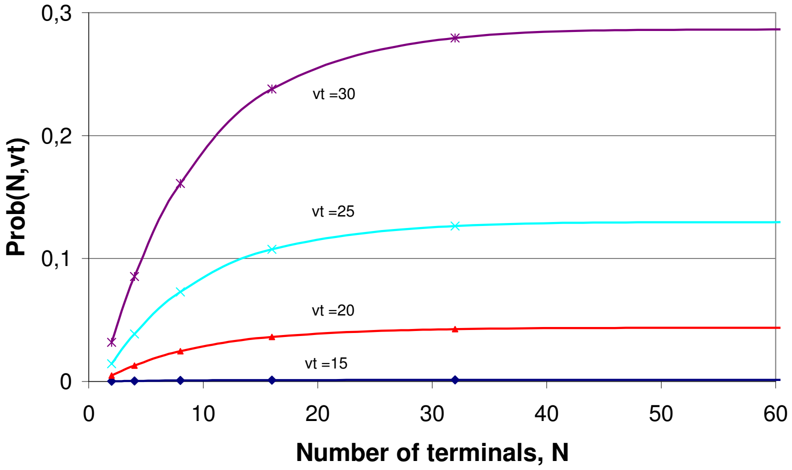


Figure 9.6: Probability of finding a usable relay mobile terminal as a function of terminals in a cell for $D = 3.5$

Figure 9.6 shows the proportion of mobile terminals available in the network which can contribute to the system performance by relaying with the fastest possible data rate. In other words, the figure depicts the likely probability that there is a mobile terminal close to the sender and moves in the direction of the destination. This probability grows slowly as the density of nodes increases in the cell. It is also shown that as the step of the mobile terminal increases, the probability increases. But increasing the step will not bring a substantial change in the result since the radius of the cellular network is relatively small as discussed in Section 3.3.1 and it is likely that the relay terminal goes out of the reach of the destination AP at once. In this sense, the proportion of terminals that takes advantage of mobility is small.

From the network model shown in Figure 9.2, the use of a mobile relay terminal is noteworthy when the sender terminal is not able to send its traffic to the AP with the higher transmission rate either directly or via a relay terminal which is positioned in its higher transmission range. Geometrically speaking, the distance between the sender and the AP should at least be greater than $2r$ for the mobile relaying to make sense. Otherwise, the problem reduces to finding the proportion of useful static relay node as in Section 5.4. The proportion of mobile relay nodes that adds to the system performance for the variation of D and number of nodes at $vt = 25$ is shown in Figure 9.7. As the number of nodes in the network increases, there is a relative increase in the probability. The sender terminal at least has a non-zero probability to transmit its traffic at a higher rate via a mobile relay terminal even if its distance from the AP has increased.

Let us take some typical values and see to what extent mobility can be useful in multi-hop cellular networks. From Figure 6.1, for a cell of radius $R = 50m$, the far terminal should get a mobile relay terminal in the range of $r = 15m$ to send its traffic at $54MBit/s$. The mobile relay terminal should at least move $vt = 20m$ to send the relayed traffic to the AP again at $54MBit/s$. That means the chance to have such a relay mobile terminal within the network is very low as shown in Figure 9.7. Alternatively, the far terminal can send its traffic with the next higher transmission rate. Equivalently, the transmission range of the terminal increases and there may not be a need to look for a mobile relay terminal as there is a higher probability to get a static relay node satisfying the condition.

Therefore, the numerical results do not optimistically justify the hypothesis that mobility increases capacity in multi-hop cellular network. But it needs further study as the analysis shown here is limited to single-relay-hop and relatively small cell size.

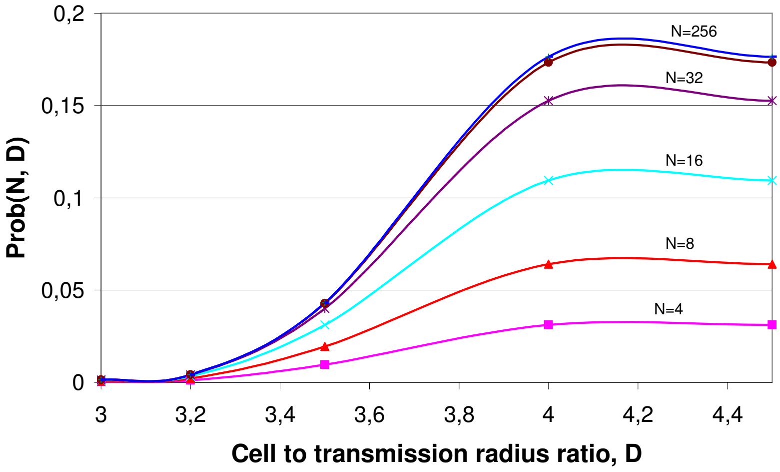


Figure 9.7: Probability of mobile terminals as a function of cell and transmission radius ratio

The numerical results shown are positive indicatives to assert the arguments given [40] that mobility improves the capacity of a network. But we should also be cautious to fully agree with such a statement as factors such as density of the terminal, velocity and delay threshold time influence the performance to a greater extent.

9.4 Simulation Setup

Consider the cellular network model treated in earlier chapters. For the static nodes, it is possible to achieve capacity gain by optimally routing and scheduling the traffic originating within the cell. Let us now introduce mobility of nodes within the cell. When terminals are allowed to move within a cell, the optimal static routing discovered may not necessarily be an optimal after a certain time. The relay terminal selected by the far terminal may not be the right relayer anymore or relay terminal that is using the second frequency assigned for relaying purpose, which is assumed to be closer to the AP, may now be at the boarder of a cell causing relatively higher co-channel interference to the neighboring cell that is using the second frequency as a base frequency. Thus, it is vital to see the extent to which the already existing static routing and scheduling is affected by the non-optimal state of the terminals in terms of capacity. It is also necessary to see the extent to which the capacity of the network is affected by increasing the rate of mobility.

For this purpose, the experiment is run on the network topology similar to the simulation set up discussed in Chapter 8. The topology consists of multiple cells in 7-cell hexagonal structure with cell radius of 50m. Similar to the static capacity evaluation case, the three different frequency layouts 3, 7, and 19 frequency reuse pattern are used.

There are n randomly distributed terminals within each cell. The terminals are either stationary or moving at an arbitrary direction depending on the speed and direction each terminal is randomly assigned as discussed in Section 9.2. When terminals reach at the boundary of a cell, they bounce at an angle depending on the incoming direction.

In this setup, every terminal is assumed to have a CBR traffic coming from an infinite source. Also the terminals are capable of relaying for upto 5 other terminals depending on their position to the AP. The system model described in Chapter 3 is used. The terminals can send their traffic either directly or via an intermediate terminal to the AP at a variable bit rate upto 54MBit/s. The performance metric used for the setup is the throughput at the AP in each cell, averaged for 55 different topologies.The metric is evaluated for direct communication, one-frequency and two-frequency relaying schemes.

9.5 Performance Evaluation

In this section the performance results are presented for the different cell layouts and relaying schemes. 55 distinct seeds are used to generate random network topology. The simulation is run for 10 steps and the performance of the network is observed at every step before the routing topology is updated and the new optimal routing and scheduling is applied.

For a pathloss coefficient $\alpha = 3.2$ and a target packet error rate of 1% Figure 9.8 shows the average throughput at the AP of a cell with 20 mobile terminals. The terminals are allowed to move within the cell at a speed $v \in [0, 10]$. The results are for one-frequency relaying mode with a uniform-time fairness scheme applied to 3, 7 and 19 frequency reuse patterns. At the 8th step the average throughput of the cell is already decreased to about 44% for 3-frequency reuse layout and about 28% for the 7- and 19- frequency reuse layout. When terminals are moving, optimal routing and scheduling algorithm becomes non-optimal in a shorter period of time and it needs frequent topology update, re-routing and re-scheduling to maintain the optimal performance. Apparently, for a slight perturbation in topology, the throughput of the 3-frequency reuse declines quickly than the 7 and 19 frequency reuse layouts. This is due to the small frequency reuse distance which is more prone to co-channel interference.

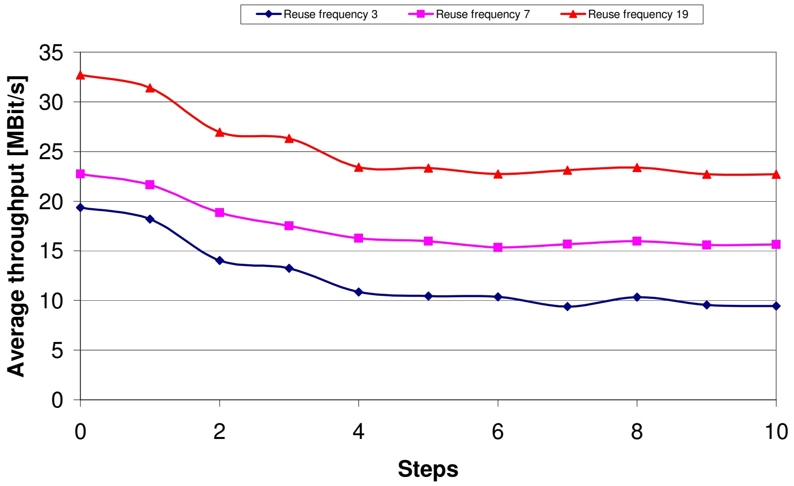


Figure 9.8: Average throughput at the AP for 20 terminals per cell

Figure **??** shows the variation of the average throughput at the AP of a cell with regular topology update. The one-frequency relaying appears more robust to topology variation than the two frequency relaying. This is due to two reasons: the optimally chosen relaying terminals will not be optimal anymore to render a higher throughput, and the optimal relaying terminals scheduled in the second frequency continue to send in same way even if they are not anymore close to the AP. This actually contributes to the additional interference in the system which eventually deteriorates the cell throughput.

Figure **??** also depicts the need for frequent update to maintain the optimal throughput and compares the results of relaying to that of direct transmission. The update frequency mainly depends on the level of throughput that can be tolerated and the computational cost due to frequent updating, routing and scheduling. For 10 steps for example, the moving average throughput shows that there is about 36% loss in throughput for two-frequency relaying. Whereas in one-frequency relaying, there is only 25% loss in throughput, yet it is about 16% superior than the two-frequency relaying. But in all the cases, relaying not only helps to improve capacity of the network, but also regulates the system performance in times of system perturbation. The throughput due to one-frequency relaying almost doubles that of direct communication and that of two-frequency relaying outperforms the direct communication by more than 66% even with sub-optimal routing information.

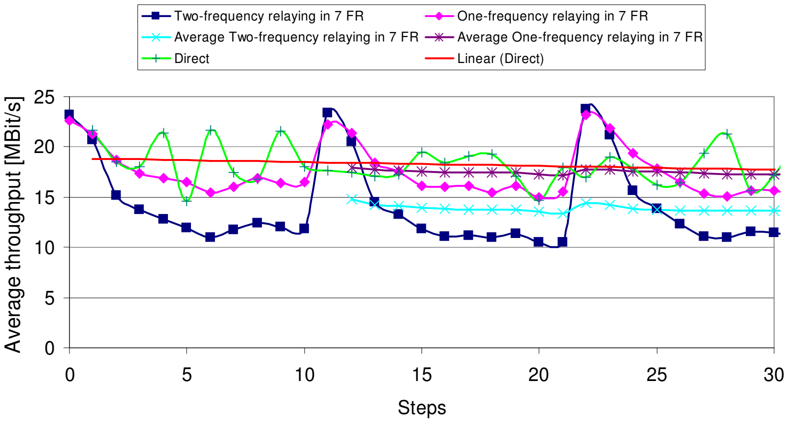


Figure 9.9: Average throughput at the AP for 20 mobile terminals with varying terminal speed

When terminals are allowed to move in different directions but with constant speed the resulting throughput variation is shown as in Figure 9.9. If terminals are moving in a selected direction at a higher speed, they may not remain within the cell. However since they are not allowed to move outside of their cell, they bounce at the boundary. If terminals repeatedly bounce at the boundary, they reduce their speed until they remain in the cell. This situation is observed in Figure 9.8, for e.g., making the throughput somewhat flat before the new update cycle comes. Also due to the bouncing effect the average throughput is fluctuating. Practically, the bouncing effect may not be manifested as terminals can keep on moving and associate themselves with a new AP.

9.6 Summary

The studies of GROSSGLAUSER and TSE [40] and BANSAL and LIU [8] suggested that use of mobility can be helpful to further improve the capacity of ad hoc wireless network. Their result does not actually suggest how the system keeps track of the dynamic network topology and regularly updates the routing to achieve a higher capacity. Taking their result as an initial hypothesis, it is attempted in this chapter to see if mobility can be useful in multi-hop cellular networks. The numerical results obtained do not substantially confirm that mobility further improves the capacity of the cellular network mainly due to the size of wireless cellular networks which are relatively small.

Furthermore, the stability of the static optimal routing and scheduling mechanisms suggested in this work are studied in the presence of mobility. It has been shown that to maintain the optimal performance of the cellular network, a frequent topology update, routing and scheduling is necessary. The results also show that the relaying mechanism suggested in this thesis have dual purpose: improving the network capacity and stabilizing the network performance if the system is unable to update the topology as frequently as needed.

Chapter 10

Comparison: Optimizing Capacity vs. Energy

10.1 Introduction

A critical design issue for wireless networks is the development of suitable communication architectures, protocols and services that increase the energy efficiency of the network and that efficiently reduce power consumption. The attempt to increase the energy efficiency has twofold advantages: the operational lifetime of mobile devices relying on limited energy stored in a battery is increased and the electromagnetic exposure of human users is potentially reduced [24, 38].

A number of different mechanisms are used to increase energy efficiency of wireless communication. One of the approaches is multi-hop communication in wireless networks. Studies on energy-efficient communication for multi-hop networks usually focus on the routing problem. They are concerned with maximizing some measures of the total energy or minimizing some function of the battery drainage. Such protocols assume that, when the physical distance of a hop is smaller, the terminals are able to appropriately reduce their transmission power, and hence, they seek to reduce a long distance hop into series of short distance [9]. The intuition why multi-hopping should improve energy efficiency in the wireless case is the non-linear relationship of distance and transmission power. The power P of a transmitted signal is attenuated over the wireless channel by a factor that is proportional to the distance d between sender and receiver and pathloss coefficient α, i.e., $P_{received} \sim P_{transmitted}/d^{\alpha}$. Hence to maintain the same signal-to-noise-ratio at the receiver, communication over shorter distances requires considerably less radiated power than over long distance. Consequently, cutting the distance in half and transmitting the same message twice is assumed to reduce the total radiated power: $d^{\alpha} > 2 \cdot (d/2)^{\alpha}$ for $\alpha > 1$. This reduction in transmission power is, however, at the cost of longer transmission time. A single message is transmitted twice (for the case of single-relay hop) instead of sending once but with a higher transmission power [25, 67].

When adaptive modulation scheme is used in the transmission, it is possible to send the messages with a faster rate and reduce the transmission time. Usually faster modulations are error-prone and may require higher transmission power to maintain the same signal-to-noise ratio at the receiver[61]. Thus, the tradeoff remains still between lower transmission power but longer transmission time and use of more complex modulation schemes that requires a better SNR, preventing the reduction of tranmission power but which allows reduced transmission power.

The core idea of this thesis lies on the use of adaptive modulation schemes for sending messages with the fastest possible rate at a reduced transmission time so as to improve the system capacity. The question now is: does multi-hopping still reduces power consumption of mobile terminals in a wireless network? This chapter evaluates the energy consumption of wireless terminals when the network capacity is optimized through multi-hopping.

The remainder of this chapter is organized as follows: The energy consumption per cell is evaluated and an example scenario is given in Section 10.2. The simulation results and discussions are presented in Section 10.3.

10.2 Evaluation of Energy

In infrastructure-based wireless network, the APs are associated with unlimited power source whereas the mobile terminals such as cellular phones, PDAs, sensors, etc., are dependent on the battery power. Thus, it is necessary to minimize their energy consumption to significantly extend the battery life time. Various energy-aware protocols and architectures have been studied and different methods of evaluating the energy consumption in wireless networks have been proposed. In this section, the energy consumption of a relaying technique which targets capacity optimization is evaluated. Energy consumption in this context is the total radiated energy per correctly transmitted bit in a cellular wireless network when terminals are sending/receiving traffic.

The energy consumption of a wireless terminal is characterized by its operational modes which are *Transmit* (TX), *Receive* (RX), *Sleep* and *Idle*, by the amplifier's power consumption and by the power needed for amplifier-unrelated parts such as baseband processing. *Idle* could refer to a terminal neither sending nor receiving, or a terminal that is powered down to a sleep mode. In this evaluation, *idle* refers to terminals neither receiving nor sending. The tightly scheduled HiperLAN/2 structure easily allows to assume sleep modes for inactive wireless terminals and hence terminals in a sleep mode are not considered in the energy consumption evaluation of the network [14, 67]. Figure 10.1 shows the values chosen for the different operation modes of the wireless terminal. P_{idle} is chosen to be considerably smaller than P_{rx}. The radiated power during transmit mode depends on factors such as distance between sender and receiver terminals, the modulation used to transmit the data, pathloss coefficient, etc., and it attains a certain maximum value for regulatory reason, in this case, for example, $200\,\text{mW}$.

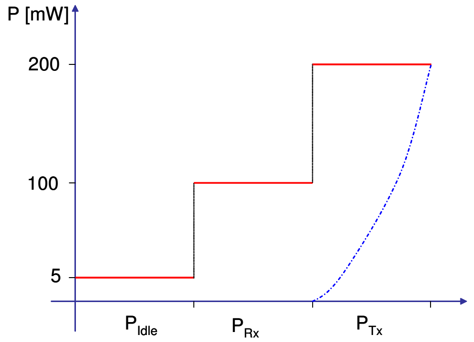


Figure 10.1: Power associated with the different operation modes of a wireless terminal

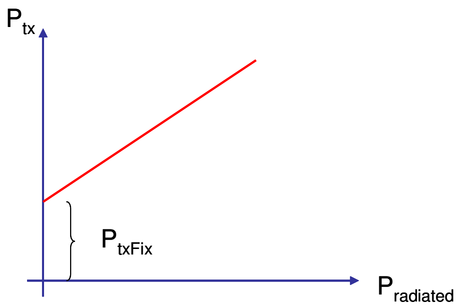


Figure 10.2: Power consumption model

For the evaluation, the power consumption model shown in Figure 10.2 is used. The model is substantiated by the measurement results in [14, 33]. It approximates the power consumption behavior of a single terminal as [67]:

$$P_{tx} = a \cdot P_{radiated} + P_{txFix} \tag{10.1}$$

where P_{tx} is the total power consumed when data is transmitted, a is a proportionality factor representing the amplifier's power consumption, $P_{radiated}$ is the network card's actual power output and P_{txFix} is the power needed for amplifier-unrelated parts such as baseband processing.

For this model, let us see how the energy consumption varies for a single terminal when sending its fixed amount of traffic to the AP with different modulations, as an example. For a certain target PER and pathloss coefficient, the optimal SNR needed to correctly receive the data is obtained for the different modulations through the mapping of target PER [61]. The radiated power $P_{radiated}$ and P_{tx} are then computed as function of distance between the sender and the receiver. Sending a fixed data with different modulation requires different amount of time. Hence the energy consumption per correctly transmitted bit associated is the product of the transmit power P_{tx} and the corresponding time need to send the data. Figure 10.3 shows the energy consumption of a single terminal for a target PER of 1% and pathloss coefficient $\alpha = 3.2$. As depicted in the figure, far terminals use slower modulation rates and relatively longer time to send a fixed amount of data, which result in higher energy consumption.

Let us now evaluate the energy consumption of the terminals for the cellular network model considered in this thesis. It has been discussed in earlier chapters, the AP is the central controller of the communication pertaining to the cell and it is responsible for computing a communication schedule for a MAC frame of 2ms long. The schedule is based on the algorithms targeting capacity optimization in the network. In a frame, each mobile terminal is assigned a time slot in which it is allowed to send or receive data to or from the access point or other mobile terminals. Moreover, this schedule also stipulates the transmission power a mobile terminal uses and one out of the seven modulation types

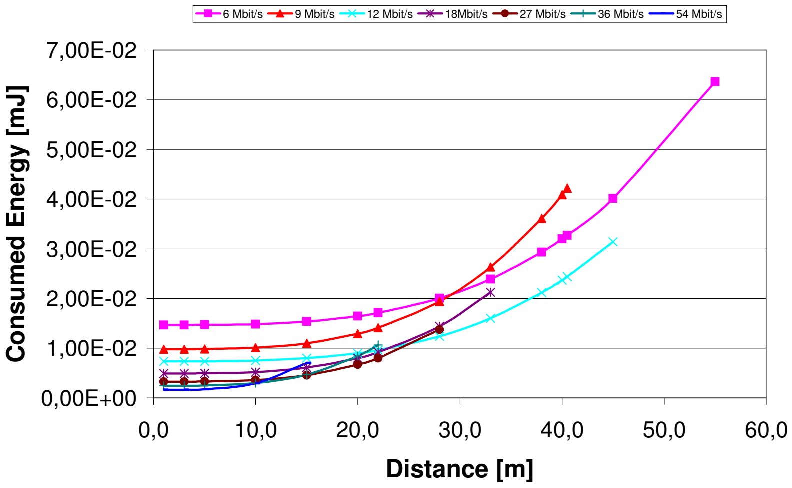


Figure 10.3: Energy consumption per correct bit for a single terminal sending a fixed amount of data to the AP

to be used within a time slot. The information needed to evaluate the energy consumption-radiated power and time slot– are thus obtained directly from the scheduled frame.

Assume that there are N terminals in a cell with data to send. The energy consumption computation for the three different schedules–direct, one-frequency and two-frequency relay communication–is as follows:

1. **Direct schedule**

   In a direct communication schedule, the terminals are assigned a time slot in a sequential manner within the communication frame. Let t_i be the time slot assigned for terminal i. The total energy consumed per correctly bit per frame with the direct schedule E_{direct} is:

$$E_{direct} = \sum_{i=1}^{N} P_{tx}^{i} \cdot t_i + P_{idle} \cdot (T_{frame} - t_i) \tag{10.2}$$

   where P_{tx}^{i} is the power consumption of terminal i as described in Equation (10.1), P_{idle} is a constant power consumed by a terminal which is neither sending nor receiving in the scheduled frame and T_{frame} is the total frame length.

2. **One-frequency relay schedule**

   In this frame schedule, terminals send their data either directly or through single-relay-hop to the AP. Let S be the set of terminals in N which use single-relay-hop. The total energy consumed per correctly bit per frame with one-frequency relay schedule $E_{relay1F}$ is:

$$E_{relay1F} = \sum_{i \in S, j \in N-S} (P_{tx}^{ij} + P_{rx}^{j}) \cdot t_{i1} + (P_{tx}^{j} \cdot t_{i2}) + P_{idle} \cdot (T_{frame} - t_{notIdle}) \\ + \sum_{k \in N-S} P_{tx}^{k} \cdot t_k + P_{idle} \cdot (T_{frame} - t_{notIdle}) \tag{10.3}$$

where P_{tx}^{ij} is the power consumption of terminal i when sending its data to the relay terminal j, P_{rx}^{j} the power needed to receive the relay data from i, t_{i1} and t_{i2} are the time needed to send the relay data from i to j and from j to the AP. Since the terminals may be assigned more than once in the schedule, the idle time is not straight forward as in the direct case. Following this, $t_{notIdle}$ should be the sum of the time slots used by terminals involved in either sending or receiving data.

3. **Two-frequency relay schedule**

   Here, two separate frames are involved at a time, one for the schedule involving the macro-frequency (the cell's primary frequency) and the other for the micro-frequency (second frequency). Terminals with relaying may be scheduled either in one- or two-frequencies depending on the number of terminals (details on scheduling relaying terminals in two-frequencies is discussed in Section 6.5). Also there may be terminals with direct communication only. Thus, this schedule is a composite of case 1 and 2 with additional frame for the second frequency. Let P be the set of terminals in N which are scheduled in two-frequency. The total energy consumed per correctly bit $E_{relay2F}$ is:

$$E_{relay2F} = \sum_{m \in P, n \in N-M-S} (P_{tx}^{mn} + P_{tx}^{n} + P_{rx}^{n}) \cdot t_m + P_{idle} \cdot (T_{frame} - t_{notIdle}) \\ + \sum_{i \in S, j \in N-S} (P_{tx}^{ij} + P_{rx}^{j}) \cdot t_{i1} + (P_{tx}^{j} \cdot t_{i2}) + P_{idle} \cdot (T_{frame} - t_{notIdle}) \\ + \sum_{k \in N-S} P_{tx}^{k} \cdot t_k + P_{idle} \cdot (T_{frame} - t_{notIdle}) \tag{10.4}$$

   where P_{tx}^{mn} is the power consumed when terminal m scheduled at the second frequency is sending to the relay terminal n and t_m is the corresponding time slot used.

Note that in all the cases the power consumption in the AP is not considered assuming that the AP is equipped with a constant power source and hence not relevant to energy efficiency.

To elaborate the above cases, consider the example scenario treated earlier in Section 6.7 with an AP and three mobile terminals in a cellular network as shown in Figure 10.4. Let the terminals M_1, M_2 and M_3 use $150, 145$ and 156mW respectively to send their data directly to the AP. Also M_1 uses

156mW to relay its data to M_2 and M_3 can only communicate directly, as indicated in the figure. The corresponding direct, one- and two-frequency relay schedules are shown in Figure 10.5.

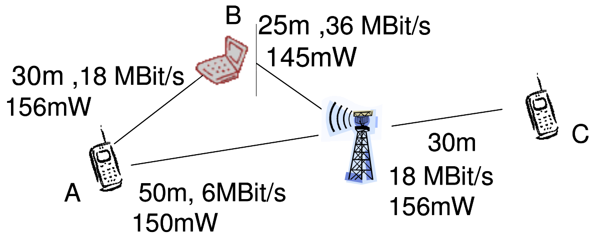


Figure 10.4: An example scenario with corresponding modulation and transmission power for a pathloss coefficient $\alpha = 3.2$

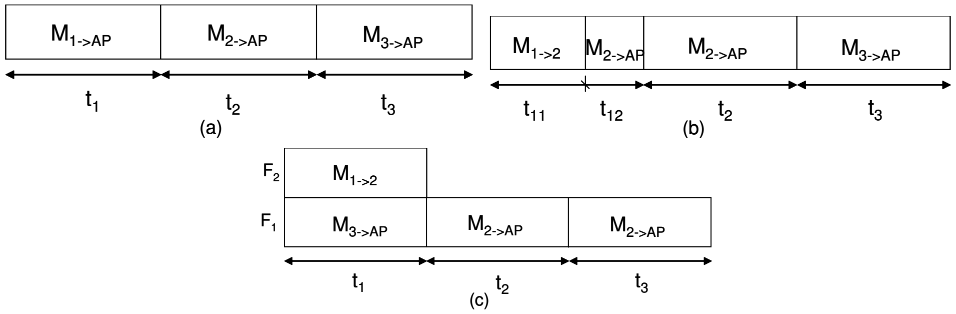


Figure 10.5: Communication schedule for (a) direct (b) one-frequency and (c) two-frequency

For the 2ms long direct frame schedule, the total energy consumed is simply $(P_{tx}^1 + P_{tx}^2 + P_{tx}^3) \cdot 2/3 + P_{idle} \cdot 4/3$. For a typical values of 5mW idle power and 100mW P_{rx} as indicated in Figure 10.1, for a constant P_{txFix} value of 50mW and for the proportionality factor $a \approx 3.45$ as in [67], the E_{direct} is 1157.3mJ. Similarly, using Equation (2) the total energy consumed per correctly transmitted bit for one-frequency schedule is $(P_{tx}^{1,2} + 100) \cdot 4/9 + P_{tx}^2 \cdot 2/9 + P_{idle} \cdot 2/3 + P_{tx}^2 \cdot 2/3 + P_{idle} \cdot 2/3 + P_{tx}^3 \cdot 2/3 + P_{idle} \cdot 4/3$, which is 1200.4mJ. And from Equation (3), the total energy consumed for two-frequency relaying schedule is 1584.6mJ.

10.3 Simulation Results and Discussion

Energy per bit is a system perspective and may not be immediately evident to the user. From a user perspective, energy consumption is relevant as it reduces the time a terminal is functional. So it is vital to see the effect of protocols and algorithms with regard to energy consumption. The previous section discusses how the energy consumption is evaluated in a multi-hop cellular network. In this section, the total energy consumption of mobile terminals in a multi-hop cellular network is presented when the network is operating at optimal capacity using the methods suggested in the thesis.

We run simulations for a multi-cellular network scenario very similar to the ones presented in Chapter 8. Figure 10.6 shows the total energy consumption of mobile terminals scheduled in the TDMA frame of the cell for a target PER of 1% and pathloss coefficient $\alpha = 3.2$. As the number of terminals in the cell increase, the total energy consumption increases correspondingly in all type of communication modes. It is interesting to note that the direct communication renders relatively smaller total energy consumption whereas the two-frequency relaying mode scores the highest.

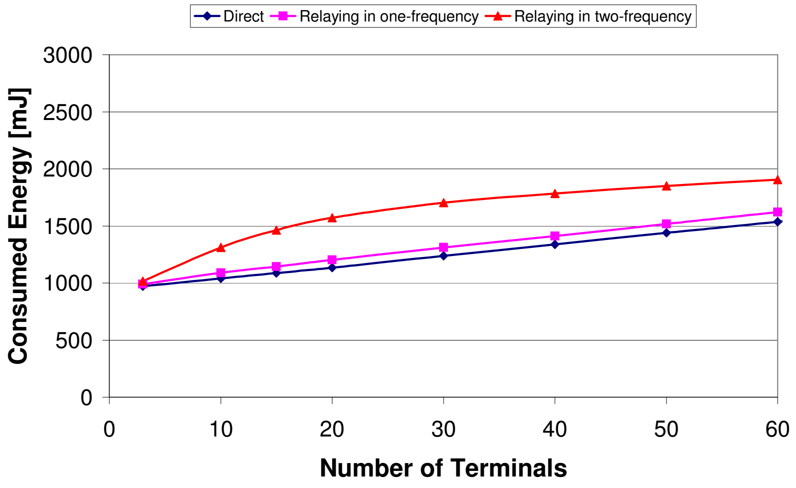


Figure 10.6: The total energy consumed in a cellular network as a function of the mobile terminals.

As indicated in Section 10.1, on the one hand, relaying reduces the communication distance and equivalently reduces the transmission power ($P_{radiated}$). On the other hand, if a faster transmission rate is needed for the reduced communication distance, higher transmission power is eventually used as faster rates are error prone. With relaying, the number of terminals participating in sending/receiving data increases, which further increases the energy consumption either in the form of $P_{receive}$ or P_{txFix}. Such facts are reflected in the result. In the direct communication mode, all terminals send their data to the AP directly and it is only the AP which is engaged in receiving mode all the time. Here, AP is assumed to have a constant power source and the power it needs to receive all the data does not account for the energy efficiency. The total energy consumption is, thus, mainly due to the radiated power to send data to the AP, depending on the distance the terminals are located from the AP and the data rate the terminals are using.

Terminals choose one- or two-frequency relaying whenever it is advantageous to have a higher throughput. This happens when there are terminals which are close to the AP and can send with fastest rate and optimal transmission power. Optimal in this sense is the smallest possible transmission power pertaining to the sender/receiver distance and corresponding modulation. An example of such relationship is shown in Figure 6.1. The point here is having shorter communication distance due to relaying may not necessarily imply a smaller transmission power and hence the increase in total energy consumption is observed. Moreover the relaying terminals involved in sending/receiving additionally contribute to the increase in the energy consumption.

Comparing the two relaying modes, one-frequency relaying is presumably scheduled for fastest possible modulation with the corresponding transmission power but with shortest time slot. From the example schedule in Figure 10.5, it can be seen that for the same amount of data the one-frequency relaying splits the time slot which otherwise is allocated for the direct mode, trading off higher transmission power against shorter transmission time. Whereas in two-frequency relaying, there is no such compromise in the transmission time slot. As a result, the two-frequency mode incurs a higher energy consumption per cell, which is approximately 23% and 18% with respect to the direct and one-frequency mode, respectively. The one-frequency relaying causes approximately a 7% increase in total energy consumption compared to the direct one. Figure 10.7 depicts the direct to relay energy consumption ratio.

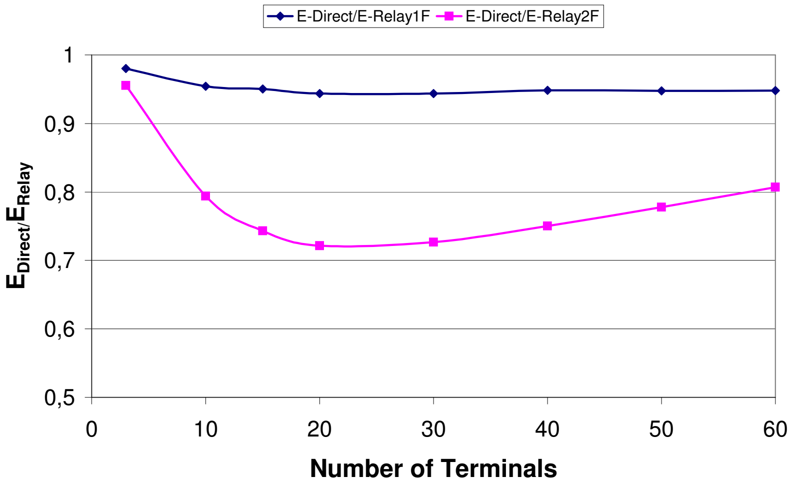


Figure 10.7: Ratio of total energy consumed in direct communication to relaying in a cellular network as a function of the mobile terminals.

Figure 10.8 shows the variation of energy consumption with respect to pathloss coefficient α in a cell 20 terminals. For smaller pathloss coefficients, communicating directly to the AP is equally advantageous as there is no difference in total energy consumption for all the communication modes.

The variation of total energy consumption due to target PER is depicted in Figure 10.9. Evidently, for higher target PER, the required signal-to-noise ratio is easily achieved without radiating high transmission power and hence the total energy consumption is relatively less. On the contrary, high transmission power is needed to achieve the small target error rate, resulting in relatively high energy consumption.

In the simulation results presented, a three-frequency reuse pattern is used. The result is more or less similar for seven- and nineteen frequency reuse pattern. This is solely due to the optimal data rate and transmission power selection algorithm which targets throughput optimization. The effect of co-channel interferences due to the different frequency reuse patters is reflected on the overall capacity

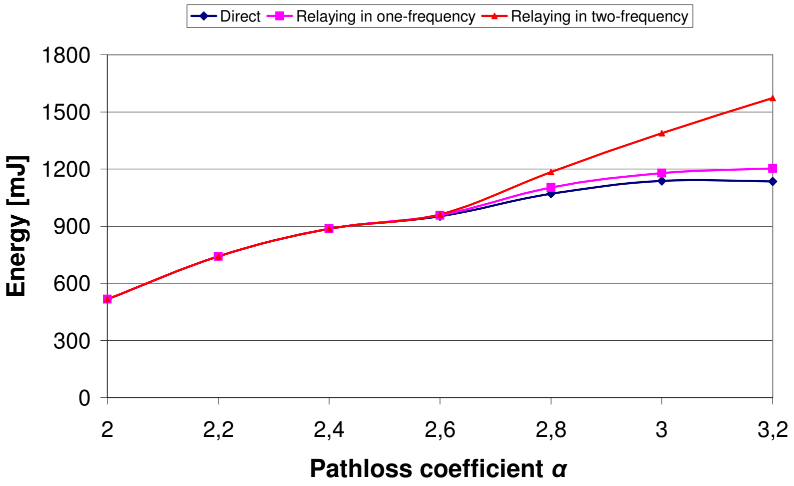


Figure 10.8: Total energy consumed by 20 mobile terminals in a cellular network for different pathloss coefficients

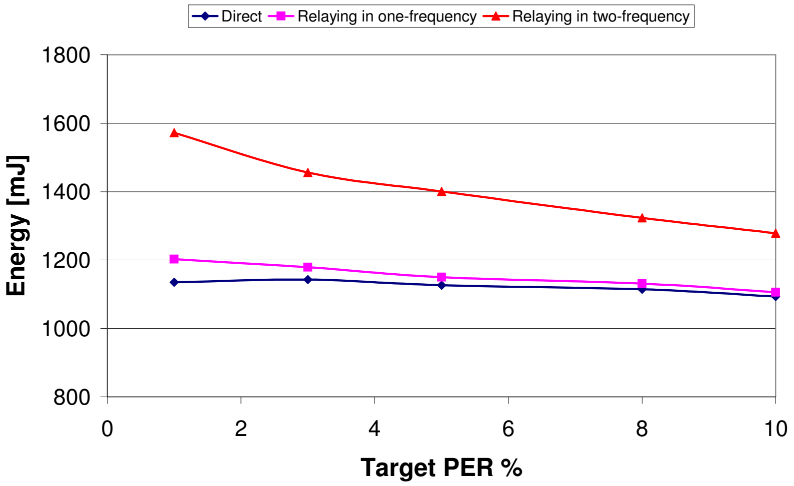


Figure 10.9: Total energy consumed by 20 mobile terminals in a cellular network for different target PER

of the network. Whenever there is high co-channel interference, the PER increases. To remain in the target PER, the signal-to-noise ratio needs to be increased by increasing the transmission power or the data rate need to be lowered until the target PER is achieved. Here, the latter is used in most cases. However, the transmission power is increased to the maximum possible limit if the communication fails to meet the target PER.

Results of relaying algorithms that target energy efficiency indicates that up to 10% energy efficiency gain is achieved due to relaying [67]. In this result, for relaying to be beneficial in the sense of energy efficiency, the sender/receiver distance should considerably be large–such conditions are rarely met for relatively small cellular networks scenario. Nevertheless, reducing the fixed power offset P_{txFix} shown in Figure 10.2 offers more possibilities for achieving a better gain.

In summary, energy consumption is a relevant metric from a user perspective as it affects the time the terminal is operational, i.e., it is an indicative of the terminal's battery lifetime. For devices such as WLAN attached to a notebook where the transmission power accounts only for small percentage of the overall power consumed by the notebook, reducing the transmission power may not significantly impact the device's operational time. In contrast, for devices such as cellular phones, PDAs, sensors, etc., reducing the transmission power may significantly extend the operational time of the device.

The total energy consumption of the cellular network using one-frequency relaying is almost comparable to that of direct communication. Relaying indeed renders a better capacity to the network without being worse in energy consumption. In the case of two-frequency relaying, it remains a trade-off between a higher capacity with slightly increased energy consumption and relatively lower energy consumption but with reduced capacity. Ultimately relaying algorithms for both capacity and energy optimization should take into account the available battery capacity and device type when choosing relay terminals.

Chapter 11

Conclusion, Open Issues and Future Work

The demand for wireless networks has increased enormously in the last few years and it is easy to anticipate that wireless local area networks will be the solution for home and office automation as they offer great flexibility and mobility. Similar to the wired counterpart, wireless networks should also satisfy user requirements such as high capacity and full connectivity among terminals. However, the capacity of a wireless network is a precious resource that cannot be arbitrarily increased. Methods of increasing the capacity are thus an essential research arena.

It has been presented that when combining the features of traditional infrastructure-based cellular networks with that of ad hoc networks, multi-hop cellular network can be formed. The potential benefits of this network are provision of larger throughput, larger system coverage, better quality of service (QoS), etc. In this thesis, relaying has been used as a viable means to improve the capacity of infrastructure-based cellular wireless network. The mechanisms used are transmission power and rate adaptive based relay selection. Relaying is chosen only when it results in better throughput, indeed, it is chosen most of the time. Also the TDMA based communication frame is fairly scheduled to obtain improved capacity.

The key technique used is to recycle a frequency from a neighboring cell to utilize it in the interior of a cell to enhance transmission parallelism, hence improving capacity by keeping the access point busy at higher data rates—the disadvantages of higher interference are outweighed by the advantages. The mechanisms are practical as they are not computationally intensive. They can be implemented as iterative online algorithms, and are based on information that can be provided by real systems. Moreover, no additional infrastructure is necessary and also the requirements on the individual terminals are quite modest. Hence, the relaying approach used can provide a simple and cheap solution to add capacity to a wireless system, particularly in highly loaded networks.

It is also interesting to see that the capacity results achieved here have a non-zero per node throughput when the node density increase. The theoretical results for wireless network shows that the maximum achievable per node throughput is $O(\frac{1}{\sqrt{n}})$ and it goes to zero for dense network [41]. Notwithstanding, it is shown that the per node throughput behaves approximately like $O(\frac{1}{\log(n)})$ and for dense network, the per node throughput remains constant.

In practical implementation, there are many points to be considered. For example, energy efficiency and reduced power consumption are important issues in wireless networks. The total energy consumed in the system is evaluated when optimizing the capacity of the network. From the results shown, the total energy consumption is increased due to relaying. There is still a tradeoff between a better capacity and battery life. Thus, energy aware relaying mechanism for improved system capacity is one possible future work.

The willingness of terminals to participate in relaying is an underlining assumption in this thesis. However, users may have low battery capacity and do not want to participate in relaying or they may not be willing at all for security reasons. There are also terminals which are apparently close to the AP and are overloaded by relaying. Such issue directs to a relay management problem which is not included here but is interesting for further study.

The frequency recycling technique presented in this thesis uses a static frequency reuse pattern for both macro- and micro-frequency reuse patterns. In practice, dynamic frequency allocation is already in use for the macro-frequency pattern. However, if load-aware dynamic macro-frequency assignment is used so that the neighboring frequency with less load is employed as the micro-frequency for relaying, further improvement can be achieved. In addition, due to the scarcity of available spectrum, it is not easy to dedicate a separate channel for relaying purpose. However, if there is a mechanism to dynamically split the already available bandwidth into a number of channels depending on the current load, and use these channels separately for the primary communication and multi-hop communication purpose, it will reduce the ensuing interferences in the system and results a better system performance. Such mechanism may require the flexibility of the system hardware. But since the hardware technology is continually improved, it cannot be impossible to have the range of flexibility.

Publications

Seble Mengesha, Holger Karl, and Adam Wolisz, ”Capacity Increase of Multi-hop Cellular WLANs Exploiting Data Rate Adaptation and Frequency Recycling”, In Proc. of MedHocNet 2004, Bodrum, Turkey, June 2004.

L. M. Feeney, D. Hollos, M. Kubisch, S. Mengesha, and H. Karl, ”A Geometric Derivation of the Probability of Finding a Relay in Multi-Rate Networks”, In Proc. of Third IFIP-TC6 Networking Conf – Networking 2004, Athens, Greece, May 2004.

S. Mengesha and H. Karl, ”Relay Routing and Scheduling for Capacity Improvement in Cellular WLANs”, In Proc. of Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt’03), Sophia-Antipolis, France, March 2003.

M. Kubisch, S. Mengesha, D. Hollos, H. Karl, and A. Wolisz, ”Applying ad-hoc relaying to improve capacity, energy efficiency, and immission in infrastructure-based WLANs”, In K. Irmscher, editor, Proc. of Kommunikation in Verteilten Systemen (KiVS), pp. 195-206, Leipzig, Germany, February 2003.

H. Karl and S. Mengesha, ”Analysing Capacity Improvements in Wireless Networks by Relaying”, In Proc. of IEEE Intl. Conf. on Wireless LANs and Home Networks, pp. 339-348, Singapore, December 2001.

S. Mengesha, H. Karl, and A. Wolisz, ”Improving Goodput by Relaying in Transmission-Power-Limited Wireless Systems”, In Proc. of Informatik 2001–31. Jahrestagung der GI, OCG, pp. 537-544, Vienna, Austria, September 2001.

H. Karl, S. Mengesha, and D. Hollos, ”Relaying in Wireless Access Networks”, Business Briefing: Wireless Technology 2002, World Markets Research Center, London, England, January 2002.

Bibliography

[1] G.N. Aggelou. *The Handbook of Ad Hoc Wireless Networks*, chapter An Integrated Platform for Ad Hoc GSM Cellular Communications. CRC Press, 2003.

[2] A. Agrawal and P.R. Kumar. Capacity Bounds for Ad Hoc and Hybrid Wireless Networks. *ACM SIGCOMM Computer Communication Review*, 34(3):71–81, 2004.

[3] A. Aguiar and J. Gross. Wireless Channel Models. Technical Report TKN-03-007, Telecommunication Networks Group, Technische Universität Berlin, April 2003.

[4] E. Altman. Capacity of Multi-service Cellular Networks with Transmission Rate Control: A Queueing Analysis. In *MOBICOM'02*, Atlanta, Georgia, USA, Sept. 2002.

[5] Y. Argyropoulos and S.P.R Kumar. Capacity improvement in cellular systems through distributed, C/I-based power control. In *IEEE Conf. on Universal Personal Communications*, volume 1, pages 164–168, Cambridge, USA, September 1996.

[6] A. Banchs and X. Pere. Distributed Weighted Fair Queuing in 802.11 Wireless LAN. In *in Proc. IEEE ICC 02*, pages 3121–3127, April 2002.

[7] A. Banchs and X. Perez. Providing throughput guarantees in IEEE 802.11 wireless LAN, in Proc. IEEE WCNC 02. In *in Proc. IEEE WCNC 02*, pages 130–138, Orlando, USA, 2002.

[8] N. Bansal and Zhen Liu. Capacity, Delay and Mobility in Wireless Ad-Hoc Networks. In *Proc. Infocom*, March 2003. `http://www.ieee-infocom.org/2003/papers/38_02.PDF`.

[9] S. Bansal, R. Gupta, R. Shorey, I. Ali, S. Razdan, and A. Misra. Energy Efficiency and Throughput for TCP Traffic in Multi-Hop Wireless Networks. In *INFOCOM*, New York, June 2002.

[10] C. Bettstetter. Smooth is Better than Sharp: A Random Mobility Model for Simulation of Wireless Networks. In *MobiCom'01*, Rom, Italy, July 2001.

[11] B. Bing, editor. *Wireless Local Area Networks*. John Wiley, Inc., New York, 2002.

[12] P. Björklund, P. Värbrand, and D. Yuan. Reousrce Optimization of Spatial TDMA in Ad Hoc Radio Networks: A Column Generation Approach. In *Proc. IEEE INFOCOM*, San Francisco, CA, March 2003.

[13] M. Bronzel, W. Rave, P. Herhold, and G. Fettweis. Interference Reduction in Single Hop Relay Networks. In *Proc. of 11th Virginia Tech Symposium on Wireless Personal Communications*, Blacksburg, VA, June 2001.

[14] B. Burns and J.-P. Ebert. Power Consumption, Throughput and Packet Error Measurements of an IEEE 802.11 WLAN Interface. Technical Report TKN-01-007, Telecommunication Networks Group, Technische Universität Berlin, August 2001.

[15] T. Camp, J. Boleng, and V. Davies. Survey of Mobility Models for Ad Hoc Network Research. *Wireless Communications and Mobile Computing(WCMC): Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2(5):483–502, 2002.

[16] I. Chlamtac, M. Conti, and J.J.-N. Liu. Mobile Ad Hoc Networking: Imperatives and Challenges. *Ad Hoc Networks*, 1(1):13–64, 2003.

[17] M. Conti. *The Handbook of Ad Hoc Wireless Networks*, chapter Body, Personal and Local Ad Hoc Wireless Networks. CRC Press, 2003.

[18] J. Crowcraft, R. Gibbens, F. Kelly, and S. Östring. Modeling Incentives for Collaboration in Mobile Ad Hoc Networks. In *WiOpt'03*, INRIA Sophia Antipolis, France, March 2003.

[19] R.L. Cruz and A.V. Santhanam. Optimal Routing, Link Scheduling and Power Control in Multi-hop Wireless Networks. In *INFOCOM*, September 2003.

[20] S.N Diggavi, M. Grossglauser, and D.N.C Tse. Even One-Dimensional Mobility Increases Ad Hoc Wireless Capacity. *IEEE Transaction on Information Theory*, August 2003.

[21] A. Doufexi, S. Armour, P. Karlsson, A. Nix, and D. Bull. A Comparison of HIPERLAN/2 and IEEE 802.11a. University of Bristol, UK.

[22] O. Dousse, F. Baccelli, and P. Thiran. Impact of Interference on Connectivity in Ad Hoc Networks.

[23] M. Dramitinos, G.D. Stamoulis, and C. Courcoubetis. Auction-based Reservation in 2.5/3G Networks. In *WiOpt'03*, INRIA Sophia Antipolis, France, March 2003.

[24] J.P Ebert, D. Hollos, H. Karl, and M. Löbbers. Does Multi-Hop Communication Reduce Electromagnetic Exposure? *The Computer Journal*, 2003.

[25] T.A. Elbatt, S.V. Krishnamurthy, D.Connors, and S.Dao. Power Management for Throughput Enhancement in Wireless Ad-Hoc Networks. In *ICC 2000*, New Orleans, LA, June 2000.

[26] N. Esseling, E. Weiss, A. Krämling, and W. Zirwas. A Multi Hop Concept for HiperLAN/2: Capacity and Interference, 2002.

[27] ETSI. Broadband Radio Access Networks (BRAN). Technical report, ETSI, TR 101 475, 2000.

[28] ETSI, France. *Broadband Radio Access Networks (BRAN) HiperLAN/2, Data Link Control (DLC)Layer*, etsi ts 101 761-1 v1.1.1 (2000-04) edition, 2000.

[29] ETSI. *Broadband Radio Access Networks (BRAN), HiperLAN/2 Physical Layer*, etsi ts 101 475 v1.1.1 (2000-4) edition, 2000.

[30] ETSI. *HiperLAN/2 Data Link Control(DLC) Layer, Broadband Radio Access Networks(BRAN)*, etsi ts 101 761-2 v1.1.1 (2000-4) edition, 2000.

[31] ETSI, France. *Broadband Radio Access Networks (BRAN), HiperLAN/2 Extension for Home Environment*, etsi ts 101 761-4 v1.1.1 (2000-06) edition, 2002.

[32] L. M. Feeney, D. Hollos, M. Kubisch, S. Mengesha, and H. Karl. A Geometric Derivation of the Probability of Finding a Relay in Multi-Rate Networks. In *Proc. of Third IFIP-TC6 Networking Conf – Networking 2004*, Athens, Greece, May 2004.

[33] L.M Feeney and M. Nilsson. Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment. In *INFOCOM'01*, Anchorage, AK, US, April 2001.

[34] M.G. Fuhl. *Smart Antenna for Second and Third Generation Mobile Communication Systems*. PhD thesis, Technische Universität Wien, March 1997.

[35] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay Tradeoff in Wireless Networks. In *Infocom 04*, volume 1, pages 464–475, March 2004.

[36] M. Gastpar and M. Vetterli. On the Capacity of Wireless Networks: The Relay Case. *IEEE*, 2002.

[37] H.-F. Geerdes and H. Karl. The Potential of Relaying in Cellular Networks. In *Proc. Intl. Network Optimization Conf.(INOC)*, Paris, France, October 2003.

[38] J. Gomez and A.T. Campbell. Conserving Transmission Power in Wireless Ad Hoc Networks. In *9th International Conference on Network Protocols*, Riverside, California, November 2001.

[39] J. Gross, I. Paoluzzi, H. Karl, and A. Wolisz. Throughput Study for a Dynamic OFDM-FDMA System with Inband Signaling. In *Proc. of Vehicular Technology Conference (VTC Spring)*, Milan, Italy, May 2004.

[40] M. Grossglauser and D. Tse. Mobility Increases the Capacity of Ad-hoc Wireless Networks. *IEEE/ACM Trans. on Networking*, 2002.

[41] P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. on Information Theory*, 1998.

[42] P. Gupta and P.R. Kumar. Towards an Information Theory of Large Networks: An Achievable Rate Region. In *IEEE Int. Symp. Info. Theory*, Washington DC, June 2001.

[43] X. Hannan, C.K Chaing, and S.K. Guan. *The Handbook of Ad Hoc Wireless Networking*, chapter Quality of Service Models for Ad Hoc Wireless Networks. CRC Press, 2003.

[44] J. He, D. Kaleshi, A. Munro, Y. Wang, A. Doufexi, J. McGeehan, and Z. Fan. Performance Investigation of IEEE 802.11 MAC in Multihop Wireless Networks. In *ACM MSWiM 2005*, Montreal, Quebec, Canada, October 2005.

[45] R. Hekmat and P.V. Mieghem. Interference in Wireless Multi-Hop Ad Hoc Networks and Its Effect on Network Capacity. *Wireless Networks*, 10(4):389–399, 2004.

[46] M. Hennhöfer, G.D Galdo, and M.Haardt. Increasing the Throughput in Wireless Multi-hop System via Spatial Multiplexing. In *Proc. of 10th WWRF meeting*, New York, USA, Oct. 2003.

[47] P. Herhold, W. Rave, and G. Fettweis. Relaying in CDMA Networks: Pathloss Reduction and Transmit Power Savings. In *VTC2003*, Korea, Spring 2003.

[48] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance Anomaly of 802.11b. In *Proc. INFOCOM*, San Francisco, CA, March 2003.

[49] G. Holland, N. Vaidya, and P. Bahl. A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks. In *Proc. 7th Ann. Intl. Conf. on Mobile Computing and Networking*, pages 236–250, Rome, Italy, 2001. ACM.

[50] D. Hollos and H. Karl. A Protocol Extension to HiperLAN/2 to Support Single-Relay Networks. In *Proc. of 1st German Workshop on Ad-Hoc Networks*, pages 91–108, Ulm, Germany, March 2002.

[51] X. Hong, M. Gerla, G. Pei, and C. Chiang. A Group Mobility Model for Ad Hoc Wireless Networks, 1999.

[52] T.-C. Hou and V. Li. Transmission Range Control in Multihop Packet Radio Networks. *IEEE Trans. on Communications*, 34:38–44, Jan. 1986.

[53] H. Hsieh and R. Sivakumar. Improving Fairness and Throughput in Multi-Hop Wireless Networks. In *ICN'01*, Colmar, France, July 2001.

[54] H. Hsieh and R. Sivakumar. Performance Comparison of Cellular and Multihop Wireless Networks: A Qualitative Study. In *SIGMETRICS/Performance*, pages 113–122, 2001.

[55] L. Jacob, Q. Qiang, R.R. Pillai, and B. Prabhakaran. MAC Protocol Enhancements and a Distributed Scheduler for QoS Guarantees over the IEEE 802.11 Wireless LANs. In *Proc. IEEE VTC-Fall*, Vancouver, Canada, September 2002.

[56] Kamal Jain, Jitendra Padhye, Venkata N. Padmanabhan, and Lili Qiu. Impact of Interference on Multi-hop Wireless Network Performance. In *MobiCom 2003*, New York, USA, September 2003. ACM Press.

[57] N. Jain, S.R. Das, and A. Nasipuri. A Multichannel CSMA MAC Protocol with Receiver-Based Channel Selection for Multihop Wireless Networks, 2001.

[58] A. Jardosh, E.M. Belding-Royer, K.C Almeroth, and S. Suri. Towards Realistic Mobility Models for Mobile Ad Hoc Networks. In *MobiCom'03*, San Diego, California,USA, September 2003.

[59] H. Karl. Ad Hoc Networks. Telecommunication Networks Group, Technische Universität Berlin, May 2003.

[60] I. Katezela and M. Naghshinen. Channel Assignment Schemes for Cellular Mobile Telecommunication Systems: A Comprehensive Survey. In *Proc. IEEE Personal Communication Magazine*, volume 82,

pages 1398–1430, 1994.

[61] J. Khun-Jush, G. Malmgren, P. Schramm, and J. Torsner. HIPERLAN Type 2 for Broadband Wireless Communication. *Ericsson Review*, 2:108–119, 2000.

[62] S. Kim, Z. Rosberg, and J. Zander. Combined Power Control and Tranmission Rate Selection in Cellular Networks. In *Proc. of the 49th IEEE Vehicular Technology Conference (Fall)*, pages 1653–1657, Amsterdam, The Netherlands, 1999.

[63] Leonard Kleinrock and John Silvester. Optimum Transmission Radii for Packet Radio Networks or Why Six is a Magic Number. *IEEE*, 1978.

[64] M. Kodialam and T. Nandagopal. Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem. In *MobiCom 2003*, New York, USA, September 2003. ACM Press.

[65] U.C. Kozat and L. Tassiulas. Throughput Capacity of Random Ad Hoc Networks with Infrastructure Support. In *MobiCom 2003*, New York, USA, September 2003. ACM Press.

[66] A. Krämling. A Power Control Strategy for HiperLAN/2. In *10th Aachen Symposium in Signal Theory (ASST)*, Aachen, Germany, September 2001.

[67] M. Kubisch, S. Mengesha, D. Hollos, H. Karl, and A. Wolisz. Applying Ad Hoc Relaying to Improve Capacity, Energy Efficiency and Immission in Infrasturcture-based WLANs. In *Proc. kivs*, Leibsig, Germany, Feb. 2003.

[68] P. Kyasanur and N. Vaidya. Capacity of Multi-Channel Wireless Networks: Impact of Number of Channels and Interfaces. In *MobiCom 05*, pages Köln, Germany, September 2005.

[69] J. Li, D. S. J. De Couto, H. I. Lee, and R. Morris. Capacity of Ad Hoc Wireless Networks. In *Proc. 7th Ann. Intl. Conf. on Mobile Computing and Networking*, pages 61–69, Rome, Italy, July 2001. ACM.

[70] J. Li, Z.J. Haas, and M. Sheng. Capacity Evaluation of Multi-Channel Multi-Hop Ad Hoc Networks. In *ICPWC02*, New Delhi, India, December 2002.

[71] J. Li, Z.J. Haas, M. Sheng, and Y. Chen. Performance Evaluation of Modified IEEE 802.11 MAC for Multi-Channel Multi-Hop Ad Hoc Network. In *Advanced Information Networking and Applications (AINA03)*, Xian, China, March 2003.

[72] Y.-D. Lin and Y.-C. Hsu. Multihop Cellular: A New Architecture for Wireless Communications. In *Proc. INFOCOM*, Tel-Aviv, Israel, March 2000.

[73] B. Liu, Z. Liu, and D. Towsley. On the Capacity of Hybrid Wireless Networks. In *Proc. IEEE INFOCOM*, San Francisco, CA, March 2003.

[74] X. Liu, E.K.P. Chong, and N. B. Shroff. Transmission Scheduling for Efficient Wireless Utilization. In *Proc. Infocom*, 2001. `http://citeseer.nj.nec.com/pdf/473896`.

[75] M. Lott and M. Weckerle. Performance Analysis of Scheduling in Wireless Networks, 2005.

[76] M. Lott, M. Weckerle, W. Zirwas, H. Li, and E. Schulz. Hierarchical Cellular Multihop Networks. In *EPMCC'03*, 2003.

[77] S. Lu, Bharghavan, and R. Srikant. Fair Scheduling in Wireless Packet Networks. *IEEE/ACM Trans. on Networking*, 7(4):473–489, 1999.

[78] H. Luo, R. Ramjee, R. Sinha, L. Li, and S. Lu. A Unified Cellular and Ad Hoc Network Architecture. In *MobiCom'03*, September 2003.

[79] P. Maille. Allowing Multi-hops in Cellular Networks: an Economic Analysis. In *ACM MSWiM'05*, Montreal, Quebec, Canada, October 2005.

[80] S. Mangold, J. Habetha, S. Choi, and C. Ngo. Co-existence and Interworking of IEEE 802.11a and ESTI BRAN HiperLAN/2 in Multi-Hop Scenarios. In *The 3rd IEEE Workshop on WLANs'01*, Boston, USA, September 2001.

[81] S. Mengesha and H. Karl. Routing and Scheduling for Capacity Improvement in WLANs. Technical Report TKN-02-012, Techniche Universität Berlin, Germany, August 2002. `http://www-tkn.ee.tu-berlin.de/publications/tknrreports.html`.

[82] S. Mengesha and H. Karl. Relay Routing and Scheduling for Capacity Improvement in Cellular WLANs. In *Proc. of WiOpt'03*, Sophia-Antipolis, France, March 2003.

[83] L.E. Milller. Probability of a Two-Hop Connection in a Random Mobile Network. In *Conf. on Information Science and Systems*. The Johns Hopkins University, March 2001.

[84] A.F. Molisch. *Wireless Communications*. John Wiley and Sons, 2005.

[85] J.P. Monks, V. Bharghavan, and W. Hwu. A Power Controlled Multiple Access Protocol for Wireless Packet Networks. In *IEEE Conf. on Computer Communications (INFOCOM)*, volume 20, pages 1–11, April 2001.

[86] J.P. Monks, J-P. Ebert, A. Wolisz, and W. W. Hwu. Energy Saving and Capacity Improvement Potential of Power Control in Multi-hop Wireless Networks. *The International Journal of Computer and Telecommunications Networking*, 41(3):313 – 330, 2003.

[87] R. Negi and A. Rajeswaran. Capacity of Ultra Wide Band Wireless Ad Hoc Networks. In *Infocom 05*, Miami, Florida, USA, March 2005.

[88] Q. Ni. Performance Analysis of Enhancements for IEEE 802.11e Wireless Networks. In *IEEE Network*, pages 21–27, July/August 2005.

[89] Q. Ni, L. Romdhani, and T. Turletti. A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN. In *Wireless Com. Mobile Computing*, pages 547–566. John Wiley and Sons, Ltd., 2004.

[90] Q. Ni and T. Turletti. QoS Support for IEEE 802.11 Wireless LAN. INRIA, Sophia Antipolis, France.

[91] T. Osche. HiperLAN/2 PHY Layer Simulation in C++. Studienarbeit, Technische Universität Berlin, Fachgebiet Telekommunikationsnetze, Berlin, Germany, July 2001.

[92] S. Papavassiliou and L. Tassiulas. Improving the Capacity in Wireless Networks Through Integrated Channel Base Station and Power Assignment. *IEEE Trans. on Vehicular Technology*, 47:417–427, 1998. `http://www.ee.umd.edu/~leandros/vehtech1.ps`.

[93] B. Penther and A. Bouttier. Performance Evaluation of HiperLAN/2 WLAN Prototype. Mitsubishi Electric Information Technology Europe, Telecommunication Lab., France.

[94] E. Perevalov and R. Blum. Delay Limited Capacity of Ad Hoc Networks: Asymptotically Optimal Transmission and Relaying Strategy. In *Infocom 03*, 2003.

[95] C. Qiao and H. Wu. iCAR: An Integrated Cellular and Ad Hoc Relay System. In *Proc. of Int'l Conference on Computer, Communications and Networks*, pages 154–161, Las Vegas, NV, USA, October 2000.

[96] D. Qiao and K. Shin. Achieving Efficient Channel Utilization and Weighted Fairness for Data Communications in IEEE. In *IEEE Int'l Workshop on QoS*, pages 227–36, 2002. `http://citeseer.ist.psu.edu/qiao02achieving.html`.

[97] J. Rapp. Increasing Throughput and QoS in a HIPERLAN/2 System with Co-Channel Interference. In *IEEE Intl. Conf. on Networking (ICN)*, pages 727–736 (Part I), Colmar, France, July 2001.

[98] T.S Rapparot. *Wireless Communications, Principles and Practices*. Prentice Hall, New Jersey, 1996.

[99] P. Samar, M.R Pearlman, and Z.J. Haas. *The Handbook of Ad Hoc Wireless Networking*, chapter Hybrid Routing: The Pursuit of an Adaptable and Scalable Routing Framework for Ad Hoc Networks. CRC Press, 2003.

[100] J. Schiller. *Mobile Communucations*. Addison-Wesley, 2000.

[101] T. J. Shepard. A Channel Access Scheme for Large Dense Packet Radio Networks. In *Proc. of Applications, Technologies, Architecture, and Protocols for Computer Communications (ACM SIGCOMM)*, pages 219–230, Palo Alto, CA, August 1996. `http://www.acm.org/pubs/citations/proceedings/comm/248156/p219-shepard/%`.

[102] A. Sheth and R. Han. Adaptive Power Control and Selective Radio Activation for Low-Power Infrastructure-Mode 802.11 LANs. In *Proc. of 23rd ICDCSW'03*. IEEE Computer Society, 2003.

[103] J. So and N. H. Vaidya. A Multi-channel MAC Protocol for Ad Hoc Wireless Networks. Technical report, University of Illinios at Urbana-Champaign, January 2003.

[104] V. Sreng. Coverage Enhancement through Two-hop Relaying in Cellular Radio Systems. Master's thesis, Carleton University, Ottawa, Ontario, 2002.

[105] V. Sreng, H. Yanikomeroglu, and D. Falconer. Coverage Enhancement through Two-hop Relaying in Cellular Radio Systems. In *Proc. IEEE Wireless Communications and Networking Conf.*, Orlando, FL, March 2002.

[106] G.L Stueber. *Principles of Mobile Communication*. Kluwer Academic Publishers, 2nd edition edition, 2001.

[107] I. Syed, M.H. Ahmed, H. Yanikomeroglu, and S. Mahmoud. Impact of Multiple Frequency Channel Usage on the Performance of TDMA-based Broadband Fixed Cellular Multihop Networks. In *WCNC'04*, Atlanta, Georgia, USA, March 2004.

[108] Z. Tang and J.J Garcia-Luna-Aceves. Hop-Reservation Multiple Access (HRMA) for Ad Hoc Networks, 1999.

[109] L. Tassiulas and S. Sarkar. Maxmin Fair Scheduling in Wireless Networks. In *Proc. IEEE 1NFO-COM'02*, 2002. `http://citeseer.ist.psu.edu/tassiulas02maxmin.html`.

[110] S. Toumpis. Capacity Bounds of Three Classes of Wireless Networks:Asymmetric, cluster and hybrid. In *MobiCom 04*, pages 133–144, May 2004.

[111] S. Toumpis and A. Goldsmith. Some Capacity Results for Ad Hoc Networks. In *Allerton Conference on Communications, Control and Computing*, volume 2, pages 775–784, Oct. 2000.

[112] S. Toumpis and A. Goldsmith. Capacity Regions for Wireless Ad Hoc Networks. In *ICC 2002*, New York, April 2002. `http://wsl.stanford.edu/Publications/Stavros/icc02.pdf`.

[113] S. Toumpis and A. Goldsmith. Capacity Regions of Wireless Ad Hoc Networks. *IEEE Transaction on Wireless Communications*, 2(4):736–748, 2003.

[114] Y-C. Tseng, S-L. Wu, C-Y. Lin, and J-P. Sheu. A Multi-Channel MAC Protocol with Power Control for Multi-Hop Mobile Ad Hoc Networks. *The Computer Journal, British Computer Society*, 45(1), 2002.

[115] A. Tzamaloukas and J.J Garcia-Luna-Aceves. Channel-Hopping Multiple Access, 2000.

[116] A. Tzamaloukas and J.J Garcia-Luna-Aceves. Receiver-Initiated Channel-Hopping for Ad Hoc Networks, 2000.

[117] N.H. Vaidya, P. Bahl, and S. Gupta. Distributed Fair Scheduling in a Wireless LAN. In *Mobile Computing and Networking*, pages 167–178, 2000. `http://citeseer.ist.psu.edu/vaidya00distributed.html`.

[118] A. Vekayutham, K. Sundaresan, and R. Sivakumar. Non-pipelined Relay Improves Throughput Performance of Wireless Ad-hoc Network. In *Infocom 05*, Miami, Florida, USA, March 2005.

[119] J. Vidal, M. Madueño, J.R. Fonollosa, S. Barbarossa, O. Gasparini, S. Ponnekanti, A. Andritsou, and A. Nix. Multihop Networks for Capacity and Coverage Enhancement in TDD/UTRAN. In *MedHoc-Net'02*, Sardegna, Italy, September 2002.

[120] B. Walke and R. Pabst. White Paper: Relay-based Deployment Concepts for Wireless and Mobile Broadband Cellular Radio. In *comnets*. rwth-aachen, Zurich, Switzerland, July 2003.

[121] L. Wang and K. Leung. A High-Capacity Cellular Network by Improved Sectorization and Interleaved Channel Assignment. In *Multiaccess, Mobility and Teletraffic for Wireless Communications, MMT'98*, volume 3, pages 43–58. Kluwer Academic Publishers, Boston, 1999.

[122] H-Y. Wei and R.D Gitlin. Two-Hop-Relay Architecture for Next Generation WWAN/WLAN Integration.

IEEE Wireless Communications, 11(2), 2004.

[123] H. Wu, C. Qiao, and O. Tonguz. Cellular and Ad hoc Relaying System. In *Proc. ICC*, volume 2, pages 450–455, 2001.

[124] H. Wu, C. Qiao, and O. Tonguz. Cellular and Ad Hoc Relaying System. In *Proc. ICC*, volume 2, pages 450–455, 2001. `http://www.ieeexplore.ieee.org/iel5/7452/20262/00936981.pdf`.

[125] S. Wu, C. Lin, Y. Tseng, and J.Sheu. A New Multi-Channel MAC Protocol with On-Demand Channel Assignment for Multi-Hop Mobile Ad Hoc Networks. In *IEEE*, 2000.

[126] H. Yanikomeroglu. Cellular Multihop Communications:Infrastructure-Based Relay Network Architecture for G4 Wireless Systems.

[127] G. Yeung, M. Takai, R. Bagrodia, A. Mehrina, and B. Daneshrad. Detailed OFDM Modeling in Network Simulation of Mobile Ad Hoc Networks. In *Proc of 18th Workshop on PADS'04*. IEEE Computer Society, 2004.

[128] S. Yi, Y.Pei, and S. Kalyanaraman. On the Capacity Improvement of Wireless Ad Hoc Networks Using Directional Antennas. In *MobiCom'03*, Annapolis, Maryland, USA, June 2003.

[129] J. Zander and S.-L Kim. *Radio Resource Management for Wireless Networks*. Artech House Publishers, 2001.

[130] H. Zhang and J.C. Hou. Capacity of Wireless Ad Hoc Networks under Ultra Wide Band with Power Constraint. In *Infocom 05*, 2005.