IMPROVING ACCESS IN NEXT-GENERATION BROADBAND WIRELESS NETWORKS

by
Paul Boone

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

School of Computer Science

at

CARLETON UNIVERSITY

Ottawa, Ontario
December, 2009

© Copyright by Paul Boone, 2009

To Mom and Dad, my wife Jenny and daughter Cate.

Abstract

As wireless technologies evolve into the next-generation, an increasing number of users will enjoy the freedom of ubiquitous broadband access. With this widespread use comes several issues that need to be addressed. We introduce and discuss our work on three areas of problem study including efficient access to data, efficient access to the network and prioritized access for emergency responders. First, we propose a hybrid XML forwarding engine for content-based routing. The forwarding engine matches stored policies against incoming XML documents. Our forwarding engine examines the type of query for matching and that has improved performance over previous works. We then propose several mobile station strategies to improve scanning, ranging and handover times in WiMAX/802.16 networks. Next, we propose two mobility profiles, based upon our scanning strategies, to further improve scanning times and handovers by adding context to MS frequency use. We evaluate our strategies and mobility profiles through simulation using both real-world mobility traces and multiple mobility patterns. Finally, we propose strategies to develop the basis of an adaptive, dual use (public and emergency), next-generation broadband wireless network. We focus on providing prioritized network access for emergency responders in mobile WiMAX/802.16 networks.

Acknowledgements

First, I must thank my wife, Jenny, who endured these long years of graduate studies. I am also grateful to and would like to thank my supervisors Dr. Evangelos Kranakis and Dr. Michel Barbeau for their support in the completion of my research. Additionally, I would like to thank my fellow students Sylvain Pitre, Michel Paquette and Christine Laurendeau who were always around to bounce ideas around or to escape for coffee. Finally, thanks is due to the Natural Sciences and Engineering Research Council of Canada (NSERC), and Mathematics of Information Technology and Complex Systems (MITACS) Network of Center of Excellence (NCE) for their support.

Table of Contents

Abstract iii

Acknowledgements iv

List of Tables xi

List of Figures xiv

List of Acronyms xvii

I Preface 1

Chapter 1 Introduction 2

1.1 Motivation 2

1.2 Summary of Contributions 5

1.2.1 Efficient Filtering of Data 5

1.2.2 Improved Network Access in WiMAX/802.16 6

1.2.3 Prioritized Emergency Network Entry Access 8

1.3 Organization of Thesis 8

II Improving Access to Data - Filtering for Content-based Routing 10

Chapter 2 Background on the Content-Based Filtering Problem 11

2.1 Introduction 11

2.2 Overview of Content-based Filtering 12

2.2.1 XPath Expression as User Subscriptions 13

2.3 Related Work 14

2.3.1 XFilter 15

2.3.2 YFilter: An NFA-based Approach Exploiting Path Sharing . . 16
2.3.3 Deterministic Finite Automata 17
2.3.4 AFilter 19
2.4 Summary 20

Chapter 3 A Hybrid XML Filtering Engine for Publish/Subscribe Content-Based Routing 21

3.1 Introduction 21
3.2 Hybrid XML Filtering Engine 21
3.2.1 DFA Structure for Simple Queries 22
3.2.2 AFilter for Complex Queries 23
3.2.3 Document Arrival and Query Matching 24
3.3 Analysis of Simple and Complex Query Structures 27
3.4 Performance Evaluation 29
3.4.1 Simulation Results 30
3.5 Conclusions and Open Problems 31

III Improving Access to WiMAX/802.16 Networks 33

Chapter 4 Overview of WiMAX/802.16 34

4.1 Introduction 34
4.2 Physical Layer 37
4.3 Medium Access Control Layer 41
4.3.1 MAC Layer Concepts 42
4.3.2 Management Connections 42
4.3.3 MAC Protocol Data Units 43
4.3.4 Bandwidth Request and Allocation 43
4.3.5 Service Flows 44
4.3.6 Quality of Service 45
4.4 Network Entry 45

4.4.1 Scanning 46
4.4.2 Ranging 47
4.4.3 Negotiating Basic Capabilities 48
4.4.4 Registration 49
4.4.5 Establishing a Service Flow 49
4.5 Mobile WiMAX and Handovers 50
4.5.1 Physical Layer 50
4.5.2 Determining Network Topology 50
4.5.3 Association Procedure 51
4.5.4 Handovers 51
4.5.5 Handover Process 52

Chapter 5 Strategies for Fast Scanning, Ranging and Handovers in WiMAX/802.16 54

5.1 Introduction 54
5.1.1 Network Entry and Handovers in WiMAX 54
5.1.2 Previous Work related to Scanning, Ranging and Handovers . 55
5.1.3 Results of the Chapter 58
5.2 Model Definition 59
5.3 Scanning Strategies 61
5.3.1 Default WiMAX Strategy 61
5.3.2 Frequency Strategies 61
5.4 Ranging Strategies 63
5.5 Previous Handover Strategy 66
5.6 A Note on MRU and MFU Strategies 67
5.7 Performance Evaluation 68
5.7.1 Simulation Setup 68
5.7.2 Simulation Results 72
5.8 Conclusion and Future Work 77
5.9 Overall Scanning Results 79

Chapter 6 Using Time-of-Day and Location-based Mobility Profiles to Improve Scanning During Handovers in WiMAX 81

6.1 Introduction . . . 81
6.1.1 Results of the Chapter . . . 82
6.2 Network Entry and Handovers in WiMAX . . . 82
6.2.1 Scanning in WiMAX . . . 83
6.2.2 WiMAX Handovers . . . 83
6.3 Revisiting Scanning Strategies . . . 84
6.3.1 Default WiMAX Strategy . . . 84
6.3.2 Frequency Scanning Strategies . . . 85
6.4 Using Mobility Profile Strategies to Improve Scanning . . . 86
6.4.1 Time-of-day Mobility Profile . . . 86
6.4.2 Location Plus Trajectory Mobility Profile . . . 87
6.5 Analysis of Manhattan Mobility Model . . . 89
6.6 Performance Evaluation . . . 90
6.6.1 Network Topology . . . 91
6.6.2 Metrics and Measurement . . . 92
6.6.3 Mobile Station Mobility Models . . . 93
6.6.4 Time-of-day Mobility Profiles . . . 94
6.6.5 Location Plus Trajectory Mobility Profiles . . . 94
6.7 Simulation Environment . . . 95
6.7.1 APRS Overview . . . 96
6.8 Simulation Results . . . 97
6.8.1 Time-of-Day Frequency Strategies . . . 97
6.8.2 Location Plus Trajectory Frequency Strategies . . . 101
6.8.3 Time-of-Day Previous Handover Strategy . . . 103
6.8.4 Mobility Trace Data Simulation . . . 107
6.9 Conclusion and Future Work . . . 109
6.10 Overall Scanning Results . . . 111

Chapter 7 Prioritized Access for Emergency Stations in Next Generation Broadband Wireless Networks 118

7.1 Introduction 118

7.1.1 Results of the Chapter 119

7.2 Background 120

7.2.1 Previous Work 122

7.3 Proposed Strategies for Prioritized Emergency Access 128

7.3.1 Emergency Contention Zone 128

7.3.2 Emergency CDMA Ranging Codes 131

7.3.3 Base Station Delayed Response 131

7.4 Analysis of Mobile Station Collisions 132

7.5 Performance Evaluation 134

7.5.1 Emergency Contention Zone 135

7.5.2 Emergency CDMA Ranging Codes 138

7.5.3 Base Station Delayed Response 139

7.6 Conclusions and Future Work 140

IV Conclusion 143

Chapter 8 Conclusions and Future Work 144

8.1 Contributions 144

8.1.1 Efficient Filtering of Data 144

8.1.2 Scanning, Ranging and Handovers in WiMAX/802.16 145

8.1.3 Mobility Profiles for Scanning and Handovers 146

8.1.4 Prioritized Emergency Access 147

8.2 Future Work 147

8.2.1 Efficient Filtering of Data 147

8.2.2 Mobility Profiles using Scanning and Handovers 148

8.2.3 Prioritized Emergency Access 148

Bibliography **150**

List of Tables

Table 3.1 XML filtering engine simulation parameters. 29

Table 3.2 XML document match times for AFilter and Hybrid structures - with 95% confidence. 30

Table 4.1 WiMAX data rates example. 39

Table 4.2 MAC layer management connections. 42

Table 5.1 Network simulation parameters. 69

Table 5.2 Proportion of time scanning can be completed with checking 5, 10 frequencies - along the path (random locations). 74

Table 5.3 Average number of ranging messages and energy expended per ranging attempt - along the path, 95% confidence. 76

Table 5.4 Average number of ranging messages and energy expended per ranging attempt - random points, 95% confidence. 76

Table 5.5 Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs, 95% confidence 77

Table 5.6 Proportion of time scanning can be completed checking between 1 and 20 frequencies - along the path, 95% confidence 79

Table 5.7 Proportion of time scanning can be completed checking between 1 and 20 frequencies - random points, 95% confidence 80

Table 6.1 Simulation parameters. 91

Table 6.2 Proportion of time scanning can be completed with checking 5, 10 frequencies - Worker class Manhattan mobility. 98

Table 6.3 Proportion of time scanning can be completed with checking 5, 10 frequencies - Worker class Manhattan mobility with detour. 99

Table 6.4 Proportion of time scanning can be completed with checking 5, 10 frequencies - Mixed classes (80%/20%) Manhattan mobility. 100

Table 6.5 Proportion of time scanning can be completed with checking 5, 10 frequencies - worker class Manhattan mobility. 102
Table 6.6 Proportion of time scanning can be completed with checking 5, 10 frequencies - Wanderer class Manhattan mobility. 103
Table 6.7 Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - Worker class Manhattan mobility, 95% confidence. 104
Table 6.8 Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - Worker class Direct mobility, 95% confidence. 105
Table 6.9 Proportion of time handover target base station chosen on first attempt - Manhattan/Direct. 105
Table 6.10 Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - Worker class Direct mobility with detour, 95% confidence. 105
Table 6.11 Proportion of time handover target base station chosen on first attempt - Mixed classes (80%/20%), Manhattan mobility. . . . 106
Table 6.12 Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - APRS mobility, 95% confidence. . . 108
Table 6.13 Proportion of time handover scanning can be completed with checking 5, 10 frequencies - APRS mobility data. 109
Table 6.14 Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Direct mobility, 95% confidence. 112
Table 6.15 Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Manhattan mobility, 95% confidence. 113
Table 6.16 Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Direct mobility with Detour, 95% confidence. 114

Table 6.17 Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Manhattan mobility, 95% confidence. 115
Table 6.18 Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Direct mobility, 95% confidence. 116
Table 6.19 Proportion of time scanning can be completed checking between 1 and 20 frequencies - APRS mobility data, 95% confidence. . . 117

Table 7.1 Simulation parameters. 135
Table 7.2 Percentage of emergency MSs processed within 50% of all MSs - split contention window. 136
Table 7.3 Percentage of emergency MSs processed within 50% of all MSs - overlapping contention window. 137
Table 7.4 Percentage of emergency MSs processed within 50% of all MSs - BS delayed response with split 50% contention window. . . . 141

List of Figures

Figure 2.1 Publish/Subscribe content based routing system. 12
Figure 2.2 An XML filtering engine. 13
Figure 2.3 An XML message and three sample XPath expressions. 14
Figure 2.4 YFilter: NFA execution. 16
Figure 2.5 Simple queries stored in a DFA. 18
Figure 2.6 AFilter's AxisView structure. 20

Figure 3.1 Simple and complex queries example. 22
Figure 3.2 Simple queries 1-4 in the DFA structure. 23
Figure 3.3 Complex queries 5-8 in the AFilter structure. 24
Figure 3.4 A sample XML document of depth 4. 25
Figure 3.5 Runtime stack for complex query matching. 26
Figure 3.6 XML document match time: Hybrid vs. AFilter. 30
Figure 3.7 Subscription structure build time: Hybrid vs. AFilter. 31
Figure 3.8 Memory usage comparison: Hybrid vs. AFilter. 32

Figure 4.1 A WiMAX/802.16 PMP network. 34
Figure 4.2 Layered architecture of WiMAX/802.16. 36
Figure 4.3 A TDD downlink subframe. 38
Figure 4.4 Model of PHY and MAC concepts. 40
Figure 4.5 The generic MAC header. 43
Figure 4.6 Network entry procedure. 46
Figure 4.7 Initial ranging process. 48
Figure 4.8 Mobile station handover from serving BS to neighboring BS 6. 52

Figure 5.1 WiMAX coverage, frequency use, four cell clusters. 60
Figure 5.2 Four cell cluster size for frequency reuse. 68
Figure 5.3 Simulation area with sample path of an MS. 70
Figure 5.4 Distribution of number of frequencies checked per scan. 73

Figure 5.5 Distribution of range request messages per ranging attempt. . 75
Figure 5.6 Previous handover strategy - BS neighbor scans per handover. 76

Figure 6.1 Mobile station handover. 84
Figure 6.2 Mobile station lookup for time-of-day mobility profile frequency ordering. 87
Figure 6.3 Overview of location plus trajectory. 88
Figure 6.4 Determining a mobile station's trajectory. 88
Figure 6.5 WiMAX coverage, frequency use, four cell clusters, Manhattan mobility model. 89
Figure 6.6 Calculating the area of a cell. 90
Figure 6.7 Simulation area with sample paths of a mobile station. 92
Figure 6.8 Frequencies checked per scan - Worker class. 98
Figure 6.9 Frequencies checked per scan - effect of a random detour on Worker class with Manhattan mobility model. 99
Figure 6.10 Frequencies checked per scan - Wanderer class and Mixed classes. 100
Figure 6.11 Frequencies checked per scan - Worker class. 101
Figure 6.12 Frequencies checked per scan - Wanderer class - LPT mobility profiles. 102
Figure 6.13 Base station neighbor scans per handover - Worker class. . . . 104
Figure 6.14 Base station neighbor scans per handover - effect of a random detour on Worker class with Manhattan mobility model. . . . 106
Figure 6.15 Base station neighbor scans per handover - Wanderer and Mixed classes. 107
Figure 6.16 Base station neighbor scans per handover - APRS mobility data. 108
Figure 6.17 Frequencies checked per scan - APRS mobility data, LPT mobility profiles. 109

Figure 7.1 Contention and data period of OFDMA frame. 120
Figure 7.2 Ranging opportunities in contention period. 121
Figure 7.3 CDMA code sent over two consecutive contention slots. 121

Figure 7.4 Split and overlapping contention windows for emergency and public MSs. 129
Figure 7.5 Threshold of arrivals of emergency MS ranging attempts and emergency contention zone. 130
Figure 7.6 Expected number of MS involved in collisions, overlapping contention window. 134
Figure 7.7 Percentage of emergency MSs processed in the first 50% of all MSs, split contention window. 136
Figure 7.8 Percentage of maximum possible emergency MSs processed in the first 50 MSs, split contention window. 137
Figure 7.9 Percentage of emergency MSs processed in the first 50% of all MSs - overlapping contention window. 138
Figure 7.10 Probability of MS collisions, overlapping contention window. . 139
Figure 7.11 Effect of assigned emergency CDMA ranging codes. 140
Figure 7.12 Percentage of emergency MSs processed in the first 50% of all MSs - split contention windows, BS delay. 140

List of Acronyms

ARQ	Automatic Repeat Request
ATM	Asynchronous Transfer Mode
AX.25	Data link layer protocol designed for use by amateur radio operators
BE	Best Effort Service
BS	Base Station
BWA	Broadband Wireless Access
CDMA	Code Division Multiple Access
CID	Connection Idenifier
CL	Candidate List
CS	Convergence Sublayer
CW	Contention Window
DCD	Downlink Channel Descriptor
DFA	Deterministic Finite Automata
DL	Downlink
DSL	Digital Subscriber Line
EIRP	Effective Isotropic Radiated Power
ertPS	Extended Real-Time Polling Service
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission

FEC	Forward Error Correction
FSM	Finite State Machine
GPS	Global Positioning System
IE	Information Element
IP	Internet Protocol
LBS	Location-Based Service
LOS	Line of Sight
MAC	Medium Access Control
MR	Multi-hop relay
MS	Mobile Station
NFA	Non-deterministic Finite Automata
NIST	National Institute of Standards and Technology
NLOS	Non Line of Sight
nrtPS	Non Real Time Polling Service
OFDM	Orthogonal Division Multiplexing
OFDMA	Orthogonal Division Multiple Access
PDU	Protocol Data Unit
PHY	Physical Layer
PMP	Point to Multipoint
QAM	Quadrature Amplitude Modulation

QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RS-BS	Relay Station-Base Station
RS	Relay Station
RSS	Received Signal Strength
RSV	Reserved
rtPS	Real Time Polling Service
SA	Security Association
SAP	Service Access Point
SDU	Service Data Unit
SFID	Service Flow Identifier
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol - acronym dropped with V1.2
SS	Subscriber Station
TDD	Time Division Duplex
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
TOD	Time-of-day
UCD	Uplink Channel Descriptor

UGS	Unsolicited Grant Service
UL	Uplink
VoIP	Voice over Internet Protocol
WiMAX	Worldwide Interoperability for Microwave Access
WL	Waiting List
XML	Extensible Markup Language
XPE	XPath Expression

Part I

Preface

Chapter 1

Introduction

1.1 Motivation

Wireless technology today has evolved to enable users to maintain continuous high-speed access to the Internet from virtually any location. This allows users to instantly access information such as email, personal files, news feeds, streaming video, and voice over IP (VoIP) services. As more and more users transition towards the freedom of access to next generation wireless networks, it will require that service providers adapt their networks to enable users to efficiently access the network and information.

With the introduction of next generation smartphones such as the iPhone and devices based on the Android or Maemo operating systems, there have been many new applications appearing that take advantage of these high-speed mobile wireless networks. Additionally, as these networks evolve, users will tend to use their mobile connections in the same manner as their at home broadband connection. In the future, users may decide to maintain a single, mobile broadband connection for all communication services, similar to the migration seen today towards adopting cellular service over a landline connection. This will make serving data to users in a timely and efficient manner as equally important to providing the network access itself.

Next generation broadband wireless networks such as WiMAX/802.16 have complex network entry procedures that require multiple steps. In order to provide better access to these networks, work needs to be done to provide faster network entry when initializing as well as during handovers. This will ensure that applications will not suffer any loss in quality of service. This is especially important during a handover where data needs to be buffered on the mobile device as well as the base station while the handover is in progress.

Since September 11, 2001, public policy has been calling for better access to communication systems for *first responders* in emergency situations [100, 92]. During

the attacks, different agencies (fire, police etc.) had difficulty communicating with each other due to the inability of their communication systems to interoperate. There is need for a broadband wireless system that allows for communication between the different agencies, as well as provide support for newer real-time applications such as telemedicine. It would be beneficial if paramedics could use portable scanning devices while transporting patients to hospital and send these images or videos ahead to the emergency room in order to get a head start on diagnosis and treatment.

This thesis will address both the access to next generation broadband wireless networks as well as access to the data traveling within the network. We have divided the work into three problems which we will now describe.

Problem 1 : Efficient Filtering of Data

As wireless networks improve in both coverage and bandwidth, allowing access to an increasing amount of information, people will expect continuous access to the Internet and instant access to information. While access to information is a powerful thing, too much information is often detrimental and causes distractions. Basex Inc., a research firm, has recently named *information overload* as "problem-of-the-year" for 2008 and suggests that the cost to the U.S. economy was $650 billion in 2006 due to lost productivity [5].

How can we efficiently filter the seemingly infinite amount of information so that users can receive only the information that is of interest to them? This leads to the publish/subscribe content-based forwarding paradigm where information "published" is only sent in the network "towards" users that have "subscribed" for it. This can result in a more efficient use of network bandwidth. Since the Extensible Markup Language (XML) has become a defacto standard for data representation on the Internet, it makes sense to focus on ways to efficiently filter XML data based on policies or rules to be defined for matching user interests. XPath is a language designed to perform queries on XML documents and is a logical choice to represent user subscriptions.

Problem 2 : Improved Network Access in WiMAX/802.16

With high bandwidth comes high demand from applications such as VoIP and streaming audio and video. These real-time applications are intolerant to delays in the network. We must ensure that the network entry procedure for a subscriber station (SS) or mobile station (MS) as well as handovers of MSs from one base station (BS) to another are completed in a timely fashion to avoid reduced quality of service (QoS).

The network entry procedure of WiMAX/802.16 is composed of several steps, as can be seen in Chapter 4. Two of the most time consuming steps are the scanning and ranging processes. One recent paper shows that the scanning process alone can be the cause of up to 90% of the delay associated with a handover for WiFi/IEEE 802.11r networks [30]. Delays during handovers can result in a degradation of quality of service (QoS) due to excessive buffering or dropped connections and it is of paramount importance to reduce these delays wherever possible.

Work needs to be done to provide fast network entry and handovers within WiMAX/802.16 networks. This can be achieved by providing more intelligent MSs that can make decisions on how to perform the scanning and ranging processes based on heuristics of past experience.

Problem 3 : Prioritized Emergency Network Access in Next-Generation Wireless Networks

Historically, wireless networks for emergency communications (i.e. police, fire) have predated and were developed independent of the public cellular networks. The presence of multiple communication systems has lead to difficulties in inter-agency communication. It is becoming a matter of public policy [100, 92] that there is a need for the development of emergency communications systems that allow for easier inter-agency communications during a time of crisis. This means that in the future, systems used by the various agencies will be required to interoperate.

Recent work from the 3rd Generation Partnership Project and others [1, 72, 75, 95]

show that there is a need for the standardization of emergency communications, especially for packet-based (data) networks. The European Telecommunications Standards Institute (ETSI) [6] is also working on standardizations for emergency communications. Also, the IEEE 802 working group recently issued a *Call For Interest* on developing a standard for emergency communications across the entire suite of IEEE 802 standards [93].

It is only recently that we have had broadband capabilities in wireless networks supporting a wide variety of multimedia applications. Perhaps now is the time to consider integrating the emergency and public networks. Work needs to be done on providing solutions for an enhanced, adaptive, dual use (public and emergency) next generation broadband wireless network. The focus will be on providing support for prioritized network access support for *authority to authority* emergency communications.

1.2 Summary of Contributions

In the following sections we will describe our contributions with respect to the three problems mentioned above.

1.2.1 Efficient Filtering of Data

The first contribution deals with Problem 1 - Efficient Filtering of Data. Users are increasingly inundated with more amounts of information and really only want information related to their interests. This brings about the content-based communication paradigm where message flow is determined by content of the message and not only a destination address. This kind of service is ideal for a publish/subscribe system. Senders (publishers) of messages are not sending their messages to any particular receiver (subscriber) and in fact may not know how many subscribers will receive each message. Subscribers use a content-based communications service to register (subscribe) to the kinds of information they are interested in, and when "someone" *publishes* data that matches their interests, or subscriptions, the data is routed by the service to the user. This changes how data is routed in a network. Rather than

simply sending data based on a destination IP address, data is forwarded towards users who have subscribed to it. This allows for more efficient use of network bandwidth as information is only sent on links towards subscribers [55]. The core of any publish/subscribe system is the forwarding/filtering engine that is responsible for matching data to rules or subscriptions.

The Extensible Markup Language (XML) [46] has become the defacto standard for data representation on the Internet. From simply providing web interfaces to databases to the Service Oriented Architecture (SOA) and SOAP, XML can be found everywhere. XPath [22] is a query language designed to search XML documents, and as such is a good choice to represent user subscriptions.

We first compiled a survey on the state of the art in the area of filtering for content-based forwarding in general as well as the concentration on XML [37]. We then described and implemented a *hybrid* XML filtering engine for publish/subscribe systems [38, 39]. Traditionally, a single structure was used to store user subscriptions when matching an XML document [27, 62, 74, 53]. Our solution differed from the traditional work since we divided the problem into different pieces by defining *types* of user subscriptions and implemented different structures to handle the different types of subscriptions. This allowed us to exploit the advantages of a deterministic finite automata (DFA) wherever possible. This XML filtering engine work was done in conjunction with the Research and Innovation Department at Alcatel-Lucent Canada, where we also filed a United States patent [44]. The work completed on the XML filtering problem is detailed in Chapter 3.

1.2.2 Improved Network Access in WiMAX/802.16

Our next set of contributions deals with Problem 2 - Improved Network Access in WiMAX/802.16. This was done by enabling more intelligent MSs that can make more informed decisions. MS decisions on how to perform steps in the network entry procedure were based on heuristics of past experience.

Improved Scanning, Ranging and Handovers

We first focused on the *scanning* phase of network entry. Scanning is performed by a mobile station (MS) in WiMAX/802.16 when it needs to gain network entry, and is the process where an MS tries to find a downlink channel from the base station (BS) by listening to the set of possible channels. This is done either on initialization of the device, or when performing a *handover* from one base station to another in order to maintain network connectivity. Previous work in this area includes [58, 83, 104] and each propose different mechanisms to improve time spent in the steps of the network entry process. We defined a number of MS based strategies to reduce the number of steps required for the scanning processes [40]. Also included in our work is the *previous handover strategy* that uses one of our scanning strategies together with the information about neighboring BSs provided to the MS by the BS. This allowed the MS to put this information into context (i.e. direction of motion) based on the results of its previous handovers.

We then extended our model to attempt to improve the time an MS spends in the *ranging* process [41]. Ranging is the next network entry step after scanning that an MS needs to perform. It is required in order to determine the correct transmit power level and timing offsets for the base station. We proposed three strategies that an MS can use to reduce the amount of time spent in the ranging process.

Mobility Profiles for Improved Scanning and Handovers

In our next work, we proposed and evaluated two mobile station mobility profiles, *time-of-day* and *location plus trajectory*, that use our scanning strategies above. The mobility profiles give a finer grained mechanism to predict the frequencies in use at the current location based on the mobility of an MS. We evaluated the performance of our mobility profiles through simulation and show further improvements in the scanning process and handovers. We simulated both with real-world mobility trace data as well as through the introduction of two simulated mobility patterns.

An additional benefit of our above strategies is that since they are MS based they require no additional network support and in fact would work well in conjunction with other previous works such as those found in [58, 78, 83, 90, 104].

We have also prepared a book chapter on an overview of WiMAX/ 802.16 broadband wireless networks [31] and it is presented in Chapter 4.

1.2.3 Prioritized Emergency Network Entry Access

The final area of research was focused on Problem 3 - Prioritized Emergency Network Access in Next-Generation Wireless Networks. We propose strategies to solve the problem in relation to developing the basis for an enhanced, adaptive, dual use (public and emergency) next generation broadband wireless network. We will focus on providing prioritized *authority to authority* communications for first responders.

Our goal was to provide seamless access to a public next generation wireless network for emergency services. We proposed a *dual use* system that gives priority to authorized emergency communications. The system operates in various states ranging from a *normal* mode with the ability to revert to an *emergency* only mode. In normal mode, all users of the system have access, with equal priority. In emergency only mode, only authorized emergency users are able to gain access to the network. The network is adaptable, initially operating in a normal mode, but having the ability to transition from to provide greater access to emergency MSs in a dynamic fashion based on currently observed communication patterns. We implemented our ideas in the context of WiMAX/802.16 networks.

In order to evaluate our system, we simulated various emergency related scenarios, from small localized events through to more serious situations leading to different loads on the system for both emergency and non-emergency communications. We will measure things such as the number of all types of connections supported, how often QoS is degraded for existing connections, and the number of preempted/blocked connections in the different modes of operation. Our work is discussed in greater detail in Chapter 7.

1.3 Organization of Thesis

The remainder of this thesis is organized as follows. Chapter 2 provides a background on content based forwarding and related works on XML filtering. Chapter 3 discusses the work on our Hybrid XML filtering engine, including architecture, simulations

and results. Chapter 4 discusses the background of the WiMAX/802.16 networks including the network entry and handover processes. Chapter 5 describes our work in the areas of improving scanning, ranging and handovers in WiMAX/802.16 networks, previous works in the area along with our performance evaluation. Chapter 6 describes two proposed mobility profiles, *time-of-day* and *location plus trajectory*, to improve the scanning operation during handovers. Chapter 7 gives examples of some previous work and describes our proposed strategies to provide prioritized network access to emergency mobile stations. Our conclusions are summarized in Chapter 8 and we discuss problems for future research.

Part II

Improving Access to Data - Filtering for Content-based Routing

Chapter 2

Background on the Content-Based Filtering Problem

2.1 Introduction

In this chapter we focus on the current state of filtering engines for publish/subscribe content-based routing systems for XML data. A publish/subscribe system is one where users (subscribers) express their interests (subscriptions) in some particular data to the system, and some service publishes data to the network [65]. The system is responsible for sending notifications to the subscribers when some data published meets the criteria of their subscriptions. In a publish/subscribe system there are two types of clients, publishers of content and subscribers of the content. Subscribers and publishers need not be aware of each other since a broker or the network will facilitate information delivery.

Some of the most advanced publish/subscribe systems are the ones used in the financial world for delivery of real-time data on stock quotes, orders, as well as corporate news [52]. These systems have now appeared on the Internet. Additional value can be brought to these systems by introducing flexibility in the ability to subscribe not only to a stock ticker symbol, but to also receive a notification when a companys stock hits a certain price, or when news about a company is published from a certain organization.

Another environment that seems well suited for content-based publish/subscribe systems is the area of sensor networks. Very large sensor networks gather enormous volumes of data that need to be aggregated somewhere. Often only a portion of the data collected is required by individual applications, for example geographic regions, or certain temperature or humidity ranges. A flexible publish/subscribe system could reduce transmission of redundant information across the network freeing up valuable resources.

A final example for the use of content-based filtering would be of help in the field

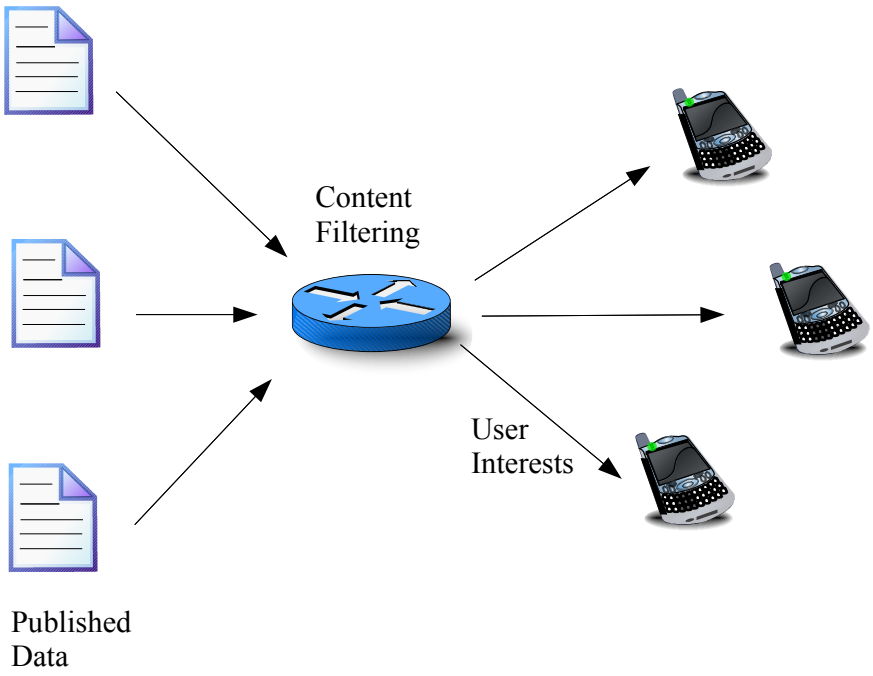


Figure 2.1: Publish/Subscribe content based routing system.

of telemedicine. In the case of ambulance services where multiple patients are arriving at multiple hospital sites, ER staff at appropriate location can receive information on the latest patient condition enroute or be notified when an incoming patient is within a few moments of arrival.

We will focus on filtering data published as XML documents against user subscriptions, represented as XPath expressions (XPEs). XML has become the defacto standard of data representation on the Web [46], and is moving towards being standardized for open document formats. XPath is a query language designed to searching XML documents [22]. The matching of XPEs can be a very complex process and is discussed in Section 2.2.1.

2.2 Overview of Content-based Filtering

At the core of any XML-based publish/subscribe system is the XML filtering engine (see Figure 2.2). The XML filtering engine is the algorithm and associated data structures used to store user subscriptions represented as XPath expressions (XPEs) and to

match published XML documents to these subscriptions. A publish/subscribe system will potentially have from thousands to millions of subscribers, each of who may have many subscriptions. The goal of the XML filtering engine is to efficiently store the user subscriptions in a structure that allows for fast matching of user subscriptions to XML documents published to the network.

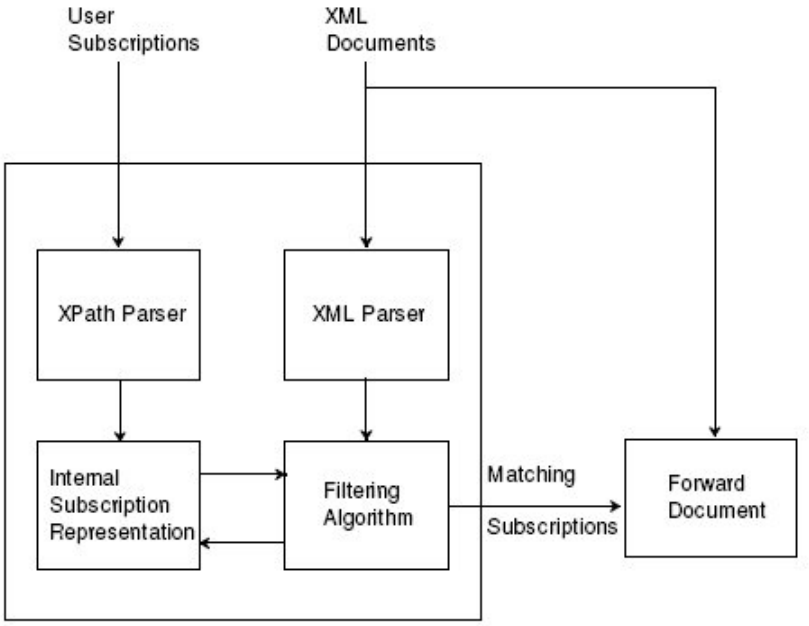


Figure 2.2: An XML filtering engine.

2.2.1 XPath Expression as User Subscriptions

There are two types of matching required in an XML filtering engine: *structure* and *value*. Structure matching deals with matching the format of the XML document itself to the subscription (query) (XPE) to be matched, while value matching deals with comparing values in location steps of XPEs to the value of a corresponding data elements in the XML document. In order to find a matching XPE for an XML document, both the document structure as well as any values present in the XPE must be matched.

Figure 2.3 shows an XML document and some XPEs representing queries that

```
<stock_quote>
    <company>
        <name>IBM</name>
        <price>$10.00</price>
    </company>
</stock_quote>

Q1 = /stock_quote/company[name="IBM"]
Q2 = /*/company/name
Q3 = //name
```

Figure 2.3: An XML message and three sample XPath expressions.

match the document. XML documents have a tree like structure. The strings between angle brackets in the XML document are called XML tags. "stock_quote" is the root tag of the XML document. Q1, Q2 and Q3 are all valid XPEs and all match the XML document. Each name in the XPath query, such as "stock_quote", "company" or "*" preceded by either a "/" or "//" is called a "location step." Query Q1 is straight forward and represents an XML document that has a root element "stock_quote" that has a child element "company" that in turn has a child element "name" that equals "IBM". Queries Q2 and Q3 are a little more complex since they introduce a wildcard, "*", and a descendent operator, "//". Wildcards and descendent operators add a level of complexity to the XML matching problem. The wildcard indicates that *any* child tag of the current open tag of the XML document will match this location step of the XPE. The descendent operator adds even more complexity since it indicates that any combination of descendants (children, grandchildren, great-grandchildren etc.) of the currently open tag of the XML document will result in a match for this location step.

2.3 Related Work

There has been a lot of work on the general content-based filtering problem with more recent work focused on matching XML documents. Previous works focus mainly on the *structure* matching of XML documents. We will focus on a subset of prior solutions that are *automata based* approaches dealing with filtering XML documents

against subscriptions represented by XPath expressions.

2.3.1 XFilter

The authors in [27] were one of the earliest automata based approaches to filtering XML documents based on XPE queries. They discuss the major components of their XFilter filtering engine for publish-subscribe system. These include (1) an event based parser for incoming XML documents; (2) an XPath parser for user profiles (subscriptions); (3) a filtering engine matches documents to profiles; and (4) their information dissemination component. Here we focus on the filtering engine. The XFilter XML filtering engine is comprised of an index structure, called Query Index, as well as modified finite state machines (FSMs). When an XML document arrives, it is parsed by the parser that sends "events" to the filtering engine handlers. Once matching profiles (subscriptions) have been matched for the document, it is then sent to the interested (subscribed) users. User subscriptions are expressed as Boolean combinations of XPath queries. XFilter converts each XPath query (user subscription) into a FSM. A subscription matches a document when the final state of its FSM is reached. Query Index is built over the states of the XPath queries. This is used to match documents to individual XPath queries.

Each XPath query is decomposed into a set of path nodes by the XPath parser. Each element node in the query is transformed into a state in the FSM. Path nodes are not generated for wildcard, "*", nodes. The Query Index is a hash table built on the element names (attributes) and contains a Candidate List (CL) and Wait List (WL) for each element. Only one node of each FSM can be active at any time and this node is called the "current node". The current nodes of all FSMs are placed in the CL of the Query Index for the nodes element name while the remaining path nodes are placed in the WL. When a state transition occurs, the new current node is promoted to the CL of its element index. The initial placement of the path nodes to the CL and WL has an impact on the performance of the system. The authors also discuss List Balancing and Prefiltering strategies to improve upon the basic system.

2.3.2 YFilter: An NFA-based Approach Exploiting Path Sharing

The authors in [62] have extended their work on XFilter. While XFilter builds a FSM for every user subscription, in YFilter, they exploit commonalities between all XPath expressions. All overlapping XPEs are merged so that common roots of XPEs are only stored and examined once. YFilter follows the same event driven philosophy of XFilter, but shares processing over the XPEs to eliminate redundant work. All user subscriptions (XPEs) are represented in a single Non-deterministic Finite Automata (NFA). The XPEs are combined into a single NFA and the transition labels form a trie over location steps of XPEs. This allows for common prefixes of paths to be represented only one time within the structure.

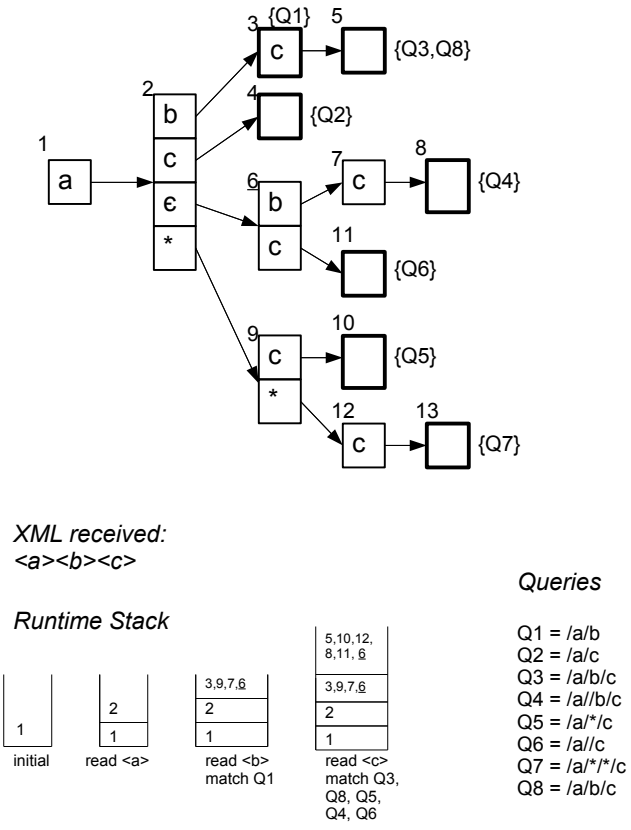


Figure 2.4: YFilter: NFA execution.

This also provided a reduction in structure memory since prefixes of overlapping queries were only stored once in the structure. YFilter improved processing time over

XFilter since multiple queries were advanced through their state machines simultaneously. However, since XPath introduces non-determinism through wildcard ("*") and descendant ("//") axes, this leads to the possibility of an exponential number of active states (since more than one state transition could occur for a given input) as you can see in the runtime stack in the Figure 2.4. Figure 2.4 shows the merged state machines of the eight queries in YFilter and a representation of the runtime stack after receiving the XML tags <a><b><c>. The currently active states are pushed on the runtime stack. Initially, the process starts at node 1, which is pushed on the stack. When <a> is received, the edge transition is made to node 2 along the link "a" and node 2 is pushed on the runtime stack. When <b> is received, the edge transition to node 3 is taken, as well as the edge transitions to nodes 6 and 9 due to the descendent and wildcard transitions from node 3. Also, since node 6 has a "b" transition and the current XML tag is "b" the transition is taken to node 7. Nodes 3, 6, 7, and 9 are pushed on the stack. Q1 has now been matched. When the next tag <c> arrives, it results in taking the transitions to nodes 5, 6, 8, 10, 11 and 12, which are pushed on the stack. Queries Q3, Q4, Q5, Q6 and Q8 have now been matched. From this simple example, it can be seen how the non-determinism caused by complex queries can result in a very large number of active states being pushed onto the runtime stack.

2.3.3 Deterministic Finite Automata

Previous to the work in [74], the use of Deterministic Finite Automata (DFA) was thought to be impractical to be used to process large numbers of XPath expressions. Typically, the number of states within the DFA explodes exponentially with the number of XPath expressions. This is due to the fact that in order to store all XPEs in a DFA, they must be "blown up" to eliminate non-determinism introduced by "*" and "//". This requires building all the possible states that might result in an XPath with "*" or "//" being matched. The authors analyze the number of states created in the DFA from XPath expressions and they found that when the DFA is created lazily and with some constraints on the structure of the XML input, the number of states does not explode. The authors approach is to simply convert the XPath query tree

into a DFA using lazy construction. A query tree may be a large collection of XPEs, and the goal is to convert them into a single DFA. This requires the construction of a NFA, then converting the NFA to a DFA. An example of a DFA can be seen in Figure 2.5 that is based on the following XPath queries:

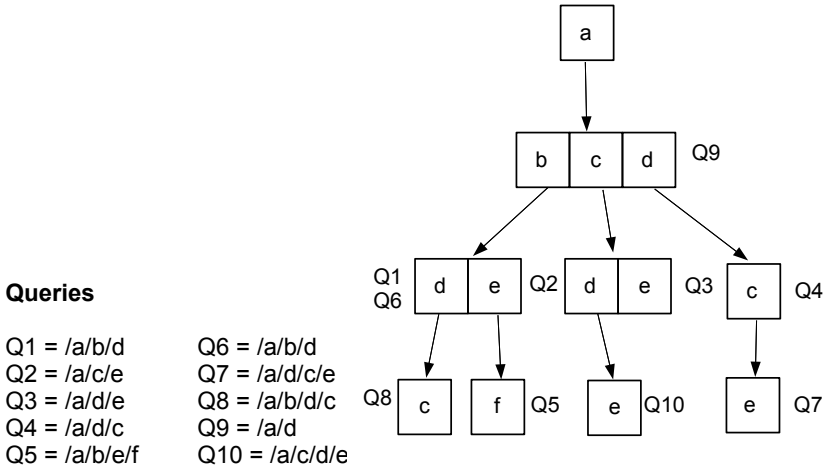


Figure 2.5: Simple queries stored in a DFA.

A lazy DFA is one that is constructed "on the fly," starting with the initial state. States are only added when an attempt is made to make a transition to a non-existing state. The hope is that at only a few of the actual DFA states will actually need to be added during XML stream parsing.

The biggest use of space in the lazy DFA is the NFA tables that need to be stored with each DFA state. If Q is the set of p XPEs of maximum depth n, then the size of each NFA table in the DFA for Q is at most n x p. From their experiments, the number of DFA states remained fairly constant for 1000, 10,000 and 100,000 XPEs. However, the size of the NFA tables grew linear to the number of XPEs.

A limitation on their work is the requirement to either know the XML document type definition (DTD) in advance, or that the recursiveness of the XML documents would be small. The recursiveness of a document refers to a structure looking like: <a><a> where the same tag can be seen as a descendent in the XML document tree. In their experiments, their filtering engine performed better than YFilter in many instances for document match times, but the

XML document structure, rather than the XPath queries impacted the results. Also, in some instances, their lazy DFA creation resulted in a large number of states and the tests could not be completed due to lack of memory.

2.3.4 AFilter

The authors of AFilter [53] proposed a new structure that is a twist on the previous NFA/DFA work. Their main structure, called AxisView, is used to represent the queries. The AxisView structure is one that (1) Only creates a single object (node) for each of the set of XML tags (observed from set of XPath queries), (2) Only puts edges between nodes that are child/descendants to parents/ancestors within queries, and (3) Keeps multiple labels for each edge with information about query identification, depth, parent/ancestor relationship, and triggering state.

The AxisView structure is shown in Figure 2.6. The edges represent transitions between location steps of queries. Labels on the edges have the form (qx, y) followed by the type of relationship between location steps. The value of x is the query number and the value of y is the depth within the query x. The relationship between location steps is that of either a parent/child ($|$) or ancestor/descendent ($\|$) and whether it is a final state ($\uparrow$) or not.

Instead of traversing states of queries as the XML document is parsed, it pushes the XML tags onto a set of stacks (StackBranch) representing the current "view" of the active tags (all tags on current XML document path that have not been closed).

In the StackBranch structure, there is a stack for each possible XML tag observed in the XPath queries. Each time a tag is received from the XML parser, a stack object is created for that element and it is pushed on its corresponding stack. For each edge in AxisView, the stack object points to the top object of the corresponding elements stack (if stack not empty). Since "/" and "//" are taken care of with labels of AxisView, that leaves the wildcard, "*" to be handled. StackBranch has a special stack dedicated for the "*" element and as elements are pushed onto their own stacks, a stack object is also created and pushed onto the "*" stack with pointers to tops of the other stacks as the "*" node edges labeled in AxisView.

The AxisView structure, shown in Figures 2.6, is only traversed when a tag arrives

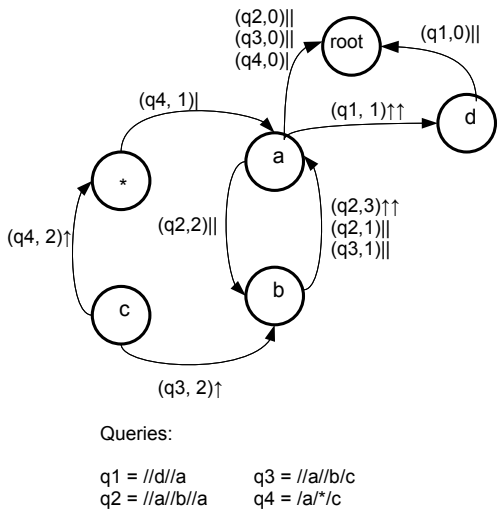


Figure 2.6: AFilter's AxisView structure.

that causes a triggering event. A triggering event is simply a tag that matches the final location step of a query (indicated by a "↑" on the AxisView label). When a triggering event occurs, AFilter uses the runtime stack (StackBranch) to traverse the AxisView structure in a bottom up (from the queries point of view) manner. If no triggering events occur, the AxisView structure is never traversed.

The authors of AFilter also described further enhancements including XPath prefix-caching and suffix-clustering. In their experiments, the AFilter solution performed better than YFilter in terms of both query matching times and memory usage.

2.4 Summary

There is still work required in improving XML document filtering. The eventual goal is to provide a filtering engine that is scalable to the Internet, requiring the storage of millions of user subscriptions and matching documents in real-time. LSI/Tarari [16] provides hardware that performs XML document matching against a database of XPEs. However, their solution is limited to 64,000 XPEs.

Chapter 3

A Hybrid XML Filtering Engine for Publish/Subscribe Content-Based Routing

3.1 Introduction

The goals of any XML filtering engine include accuracy, fast real-time XML message processing, minimal structure maintenance times and minimal memory requirements. When a published XML document arrives in the filtering engine, it must be evaluated against the set of user subscriptions in real-time. The result of the operation will produce a list of *all* subscriptions that the document matched so that it may be forwarded to the subscribers.

The current research on XML filtering focuses on designing a single "catch-all" structure to store and index the set of XPath expressions that represent subscriptions. In this chapter, we develop a new *hybrid* strategy for building an XML filtering engine and evaluate its implementation. Our Hybrid XML filtering engine breaks from tradition by examining the type of XPath query and then storing the subscriptions of different "types" in appropriate structures.

Ideally, one would prefer to use a DFA in order to store the states (location steps) of queries since this would result in the most efficient matching times. This however may not be possible since even a lazy DFA can run into memory problems due to the nature of the XPath language [74]. We introduce our Hybrid filtering engine that can exploit the advantages of a DFA wherever possible.

3.2 Hybrid XML Filtering Engine

In previous works, it is often discussed that the probability of wildcards, "*", and descendants, "//", are low. Test parameters for the wildcard and descendants probabilities typically fall in the range of 0.1 on the low end to a high of 0.3. If a high

proportion of XPath queries do not contain these non-deterministic elements, perhaps we can take advantage of that fact.

In this section we introduce our Hybrid XML Filtering Engine. This strategy looks at the "query mix" (types of XPath expression) and divides them into two types. The first type is a *simple* query that contains no descendants or wildcards. The second query type is a *complex* query that contains descendants or wildcards.

Depending on the query type the query is stored in the appropriate structure. We will store simple XPath queries in a combined DFA structure, while using the AFilter structure to store the more complex queries. As the filtering engine receives XML tags from the document parser, filtering can be performed in both structures in parallel. Figure 3.1 shows eight XPath queries consisting of both *simple* and *complex* types.

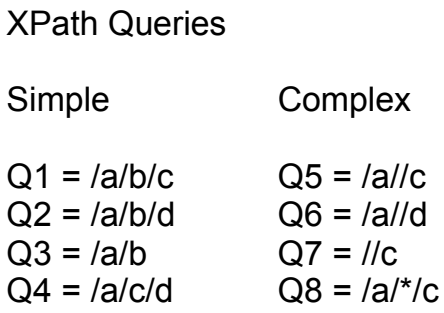


Figure 3.1: Simple and complex queries example.

3.2.1 DFA Structure for Simple Queries

Simple XPath queries, ones without wildcards or descendants, will be stored directly in a combined DFA. From [62, 74] we see that we can exploit commonalities among the XPath expressions. By using a combined DFA, common prefixes of the XPath query paths are merged and therefore are only processed once. Since this is a DFA structure, we know that we will only traverse at most one edge/transition in the DFA for each tag processed in the XML document. Furthermore, the instant we cannot

make a transition to a state associated with a received tag (non-existing state) we know that further matches along the current document path are not possible and can ignore incoming tag event processing in the DFA until such a time we receive the current states end tag.

Figure 3.2 shows how the simple queries 1 through 4 are stored in the DFA-based simple query data structure.

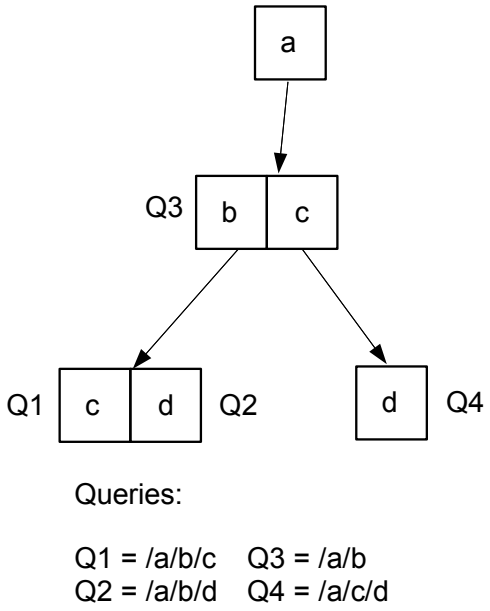


Figure 3.2: Simple queries 1-4 in the DFA structure.

3.2.2 AFilter for Complex Queries

This structure will store the more complex XPath queries. One of the benefits of the AFilter structure is that it does not traverse states like with the DFA. Since we are trying to match the more complex queries where either (1) any ("*") location step or (2) "any number of" ("//") location steps between a parent/ancestor and child/descendant, traversing an unknown expected number of states is wasteful. The AFilter structure allows us to simply check if the current tag to be processed from the XML document is the final (triggering) state (location step) of an XPath query before trying to traverse the set of query states in search of an XML document structural

match. If an XML document does not contain a triggering state, the AxisView reverse-state-machine will never be traversed.

Figure 3.3 shows how the complex queries from above would be stored in the AFilter complex query data structure.

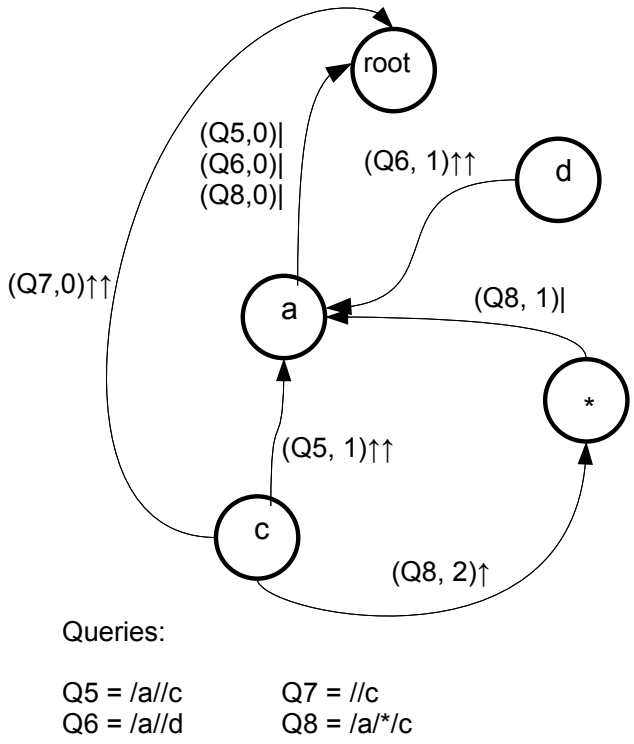


Figure 3.3: Complex queries 5-8 in the AFilter structure.

3.2.3 Document Arrival and Query Matching

Based upon the queries Q1 through Q8 that are stored in our Hybrid XML forwarding engine assume the XML document shown in Figure 3.4 arrives at the router. Here we give a step by step account of the internal workings of our Hybrid XML filtering engine. The XML document tags will arrive as follows:
<a><c></c><d><e></e></d>.

In the simple query structure, we maintain a pointer that indicates the last XML tag we have seen that matches with one or more simple queries. This is initially set to the root node of the DFA.

```
<a>
    <b>
        <c></c>
        <d>
            <e></e>
        </d>
    </b>
</a>
```

Figure 3.4: A sample XML document of depth 4.

1. The arrival of the <a> XML tag triggers the operation of the filtering engine by searching both the simple query and complex query structures in parallel. There are no simple queries matched in this step. The location pointer of the simple query structure will be updated to follow the "a" link from the root. In the case of the complex query structure, the <a> tag will be pushed on the "main" and "wildcard" stacks and it will look for any queries that have "a" as a final step indicating a triggering state (↑↑). There are none.

2. Next the <b> tag arrives. In the simple query structure, the location pointer now follows the link from "b". The simple query q3 have been matched. For the complex query structure, the "b" tag is pushed on the stacks. It then examines all the out edges from "b" in the AxisView structure and creates corresponding links from the stack object "b" to stack objects below it. No links are added from "b" in the document path tags stack, but a link from the "b" object in the wildcard stack to "a" is added since in the complex structure there is a link from "*" to "a". This enables the stack to maintain a mechanism to traverse and match queries when triggering states are reached. No triggering locations have been hit.

3. Next the <c> tag arrives. In the simple query structure Q1 is matched. There are no lower subtrees found so no further work is required in this structure until the </c> tag is received.

   Figure 3.5 shows the resulting runtime stacks. The "c" tag object is pushed on the stacks. Links are added to "a" and "q_root" from the "c" object in the

document tag stack and to “a” from the “c” object in the wildcard stack.

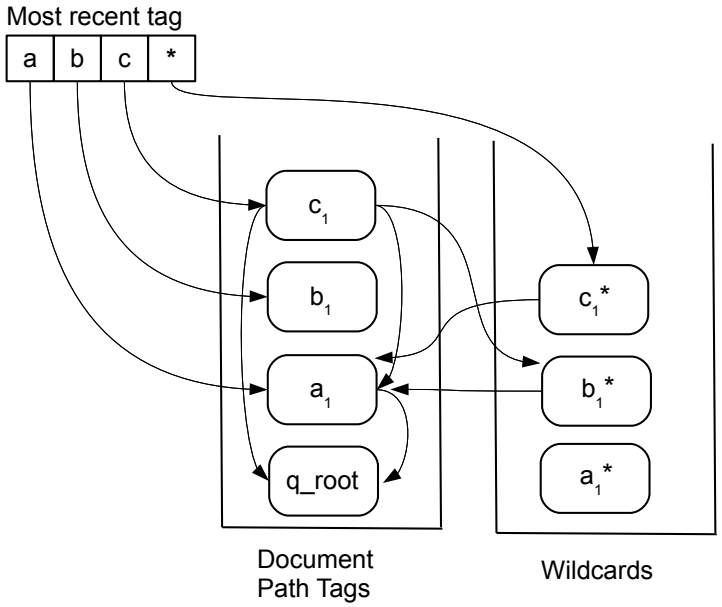


Figure 3.5: Runtime stack for complex query matching.

4. Next the </c> tag arrives. With the arrival of the closing tag for “c”, we must represent a move up a level inside the XML document. In the simple query structure, this is simply indicating that the closing tag was received and it will process the next incoming tags.

   Now queries Q5, Q7 and Q8 have been triggered at state “c” in the complex structure. The algorithm is for each query to traverse all routes via pointers within the stacks until no further location steps of the query are found or the “q_root” node is reached. The “q_root” node can be reached via the “a” node and therefore Q5 has been matched. The “c” stack object has a direct link to “q_root” therefore Q7 is matched. Q8 is also matched since “q_root” can be reached following location steps through “b” and “a”. The “c” stack objects are popped from the stacks.

5. Next the <d> tag arrives. In the simple query structure Q2 is matched. There are no lower subtrees found so no further work is required in this structure until

the </d> tag is received.

The "d" tag object is pushed on the stacks. A link is added to "a" from the "c" object in the document tag stack and to "a" from the "c" object in the wildcard stack.

6. Next the <e> tag arrives. Only the complex structure is updated. The "e" object is pushed on both stacks. There are no links that need to be updated since there are no "e" elements in queries stored.

7. Next the </e> tag arrives. No queries are matched. The "e" objects are popped from the runtime stacks of the complex structure.

8. Next the </d> tag arrives. In the simple query structure it is indicated that the closing tag was received and it will process the next incoming tags. In the complex query structure, Q6 has hit a triggering state. All routes via pointers within the stacks are traversed until no further location steps of Q6 are found or the "q_root" node is reached. The "q_root" node can be reached via the "a" node and therefore Q6 has been matched. The "d" stack objects are popped from the stacks.

9. Next the tag arrives. In the simple query structure, the location pointer is now updated to the parent node containing "b" and "c". In the complex query structure, the "b" stack object are popped from the stacks. No queries are matched.

10. Next the tag arrives. In the simple query structure, the location pointer is now updated to the parent node which is the root. In the complex query structure, the "a" stack object are popped from the stacks. No queries are matched.

3.3 Analysis of Simple and Complex Query Structures

In this section we discuss the differences in operational runtime complexity between the simple DFA and complex AFilter structures when matching the structure of XML documents against stored queries.

The number of nodes in the DFA structure is based upon the number of XML tags queries and the XML schema. However, the DFA stores identical prefixes of queries only once. That is, adding queries $/a/b/c$ and $/a/b/d$ would only add nodes a, b, c, and d to the DFA structure. The list of queries terminating at a node are maintained. There are no additional labels on edges required.

The number of nodes inside the AFilter AxisView structure is less than that of the DFA as it is determined by the number of unique XML tags found in the set of queries (a subset of the total number of unique tags found in the XML schema). For example, the tag "c" is only stored once for all queries that has "c" as a "location step". In order to maintain the queries, each edge between nodes has a label for *every* query that has a "location step" between the two nodes. The label indicates the query ID, which location step, whether it is a "/" or "//" relationship and if the edge is the final "location step" of the query.

Upon the arrival of an XML tag the operation of each structure takes at most:

1. DFA

   $O(1)$ time. The current location pointer within the structure will at most move down or up one level within the DFA. The list of queries matched is noted at each node. Additionally, the DFA does nothing once it has received an XML start tag (i.e. <a>) for which it has no direct link from its current location until the corresponding end tag (i.e.) has been received.

2. AFilter

   O(depth of stack $\times$ max # labels) in the case that the arrival triggers on all queries on all edges from the tag stack object.

Since the runtime operation of processing an XML tag of the DFA is $O(1)$, this results in processing an XML document in linear time. If a large percentage of queries stored are of the simple type, then an XML filtering engine can benefit from storing simple queries in a

3.4 Performance Evaluation

A series of experiments were performed with a varying number of stored XPath subscriptions to compare the performance of our Hybrid filtering engine against the AFilter filtering engine. Each filtering engine was populated with identical sets of XPath subscriptions and the same set of XML documents were matched against the stored queries in both filtering engines. Several key measures were collected:

1. XML document match time: the time required to determine all queries matched by a given document;

2. XML filter construct time: the time required to build the structure(s) containing XPath queries; and

3. XML filtering engine memory requirements.

Simulation Parameters	
Number of queries	10k - 60k
Depth of XML docs and queries	3 - 10
Complex query probability	10%
XML file size	1 - 8kb
Number of XML documents	280 (35 of each depth)

Table 3.1: XML filtering engine simulation parameters.

Tests will vary the following parameters, as shown in 3.1 to determine the scalability of the implementation in terms of time and memory. The query mix was 10% complex queries. For the XPath queries that were complex, the probability of a wildcard (“*”) or descendant (“//”) at any given location step was 0.3, with the “*” or “//” being chosen equally likely with probability 0.5. XML documents were of depth 3 to 10 and ranged from 1kb to 8kb in size. XPath expression queries stored in the filter were of depth of 3 to 10 location steps. The XML documents and XPath queries used were randomly generated based upon an arbitrary XML document type definition (DTD).

3.4.1 Simulation Results

Figure 3.6 shows the Hybrid engine vs. the AFilter engine XML document matching times for the parameters shown in Table 3.1. The Hybrid engine showed between 23% and 32% improvement over the AFilter for the parameters tested. The results are summarized in Table 3.2.

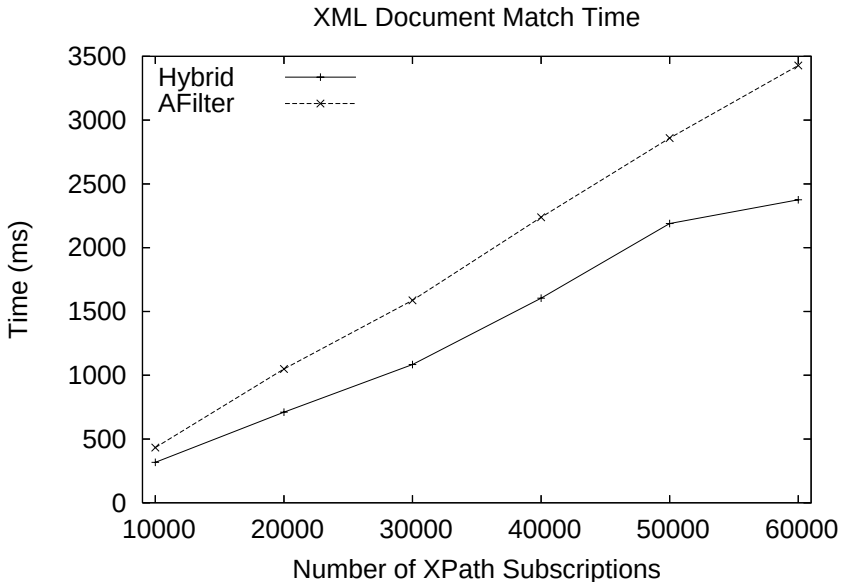


Figure 3.6: XML document match time: Hybrid vs. AFilter.

XML Document Match Time			
No. XPEs	**Hybrid**	**AFilter**	**Improvement**
10000	317.31 ± 26.72 ms	432.68 ± 36.24 ms	26.66%
20000	710.23 ± 60.19 ms	1049.34 ± 93.58 ms	32.32%
30000	1084.05 ± 94.05 ms	1586.92 ± 138.84 ms	31.69%
40000	1604.22 ± 140.28 ms	2238.67 ± 193.96 ms	28.34%
50000	2189.45 ± 195.33 ms	2859.47 ± 244.85 ms	23.43%
60000	2375.67 ± 201.83 ms	3428.90 ± 298.39 ms	30.72%

Table 3.2: XML document match times for AFilter and Hybrid structures - with 95% confidence.

A series of simulations were run to determine the XML Filter build times. Figure 3.7 shows the time to build the internal XPath query structures for the Hybrid and

AFilter filters. The build times have been reduced by between 13%-18% in the Hybrid filter. It is of interest to note that the time required to build the AFilter structure grows at a faster rate than that of our proposed Hybrid filtering engine as the number of stored queries increases.

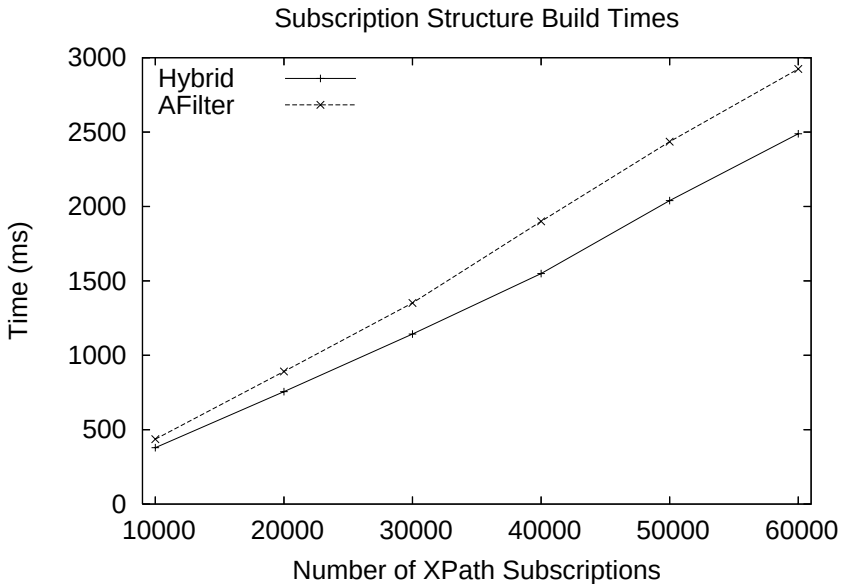


Figure 3.7: Subscription structure build time: Hybrid vs. AFilter.

Finally, a measurement of the memory used to store the XPath queries was taken for a series of simulations. Figure 3.8 shows the memory usage of the Hybrid and AFilter engines. For this test, we measured the memory use for 10,000, 50,000 and 100,000 XPath queries. The memory usage has been reduced between 12%-18% in the Hybrid filter. As with the structure build times, it is interesting to note that the memory requirements of AFilter grows at a faster rate than that of our proposed Hybrid filtering engine as the number of stored queries increases.

3.5 Conclusions and Open Problems

In this chapter we have introduced our Hybrid XML filtering engine for content-based routing. Our filtering engine examines the type of XPath queries and stores queries of different types in different structures. We compared the performance of our filtering

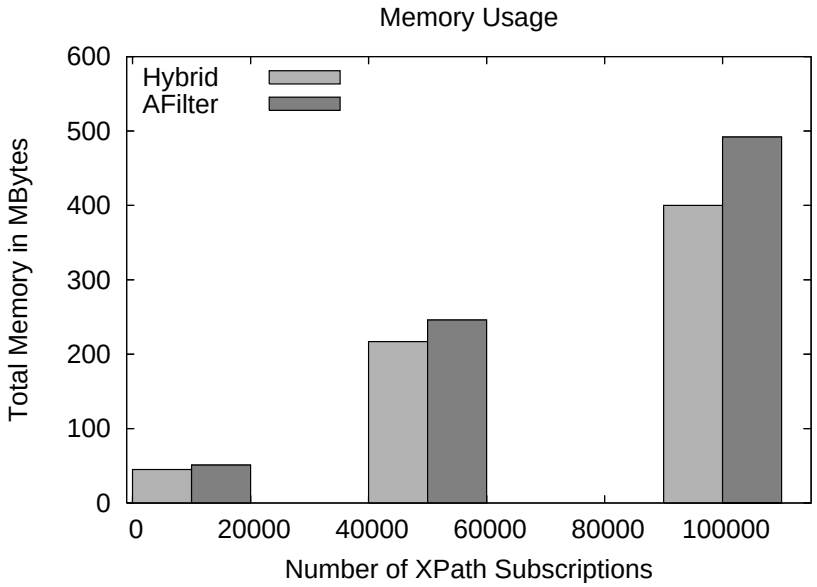


Figure 3.8: Memory usage comparison: Hybrid vs. AFilter.

engine with that of AFilter and observed improvements in all of our key measurements of XML match time, structure build time and total memory used.

There are still a number of open problems that remain. Implementation of prefix-caching and suffix-clustering would further improve matching times [53]. Some empirical studies on the actual "query mix" for XPath queries for publish/subscribe systems should be performed in order to better understand their nature.

There may be other strategies that might help in improving filtering performance. Maybe the complex queries can be further decomposed into ones with wild cards, ones with descendent operators or ones with both. Perhaps we can store XPath queries based on their depth.

The implementation of the AFilter filtering engine used for comparison as well as to store complex queries in our Hybrid filtering engine was not optimized and thus had less than the expected performance. Finally, the main focus of most previous and current work is on *structure* matching, the optimization of *value* matching needs to be investigated in the context of an XML matching engine. There may be many queries that have overlapping prefixes, resulting in many different values that must be compared at each common location step.

Part III

Improving Access to WiMAX/802.16 Networks

Chapter 4

Overview of WiMAX/802.16

4.1 Introduction

Worldwide Interoperability for Microwave Access/IEEE 802.16 (WiMAX/802.16) is a technology for fixed or mobile and secure broadband wireless access (BWA). It is a wireless alternative to digital subscriber line (DSL) technology. The main envisioned applications are wireless voice over IP, wireless Internet access and broadband wireless access in rural areas. Topics covered in this chapter include an overview of WiMAX/802.16, the physical (PHY) and medium access control (MAC) layers, mobility, mesh mode and multihop relay.

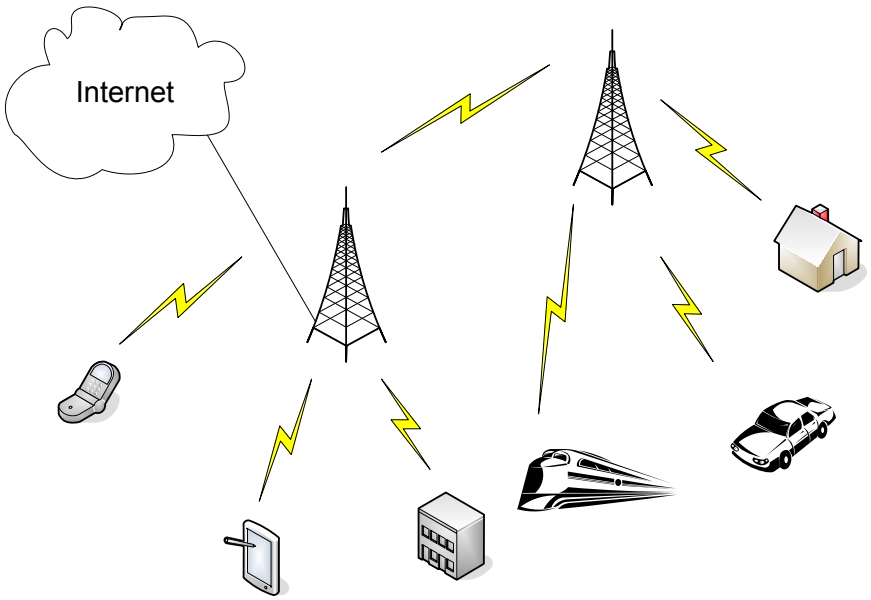


Figure 4.1: A WiMAX/802.16 PMP network.

WiMAX/802.16 was first published as the IEEE 802.16 standard that defines fixed point-to-multipoint (PMP) BWA in the 10 to 66 gigahertz (GHz) range. Amendment 802.16a adds control enhancements, mesh mode support, more frequencies in the 2 to 11 GHz range, non-line-of sight communications and licensed or unlicensed operation. Amendment 802.16c specifies implementation profiles. This work has been combined into the 802.16-2004 standard [10]. Mobility support and improved security are introduced in 802.16e-2005 [9]. A management information base (MIB) is documented in 802.16f. Management plane procedures and services are covered in 802.16g. At this time, the latest full standard has been incorporated as 802.16-2008. The WiMAX Forum is an organization that addresses implementation certification of WiMAX/802.16 products and integration with second and third generation cellular technologies.

The main mode of operation of WiMAX/802.16 is point-to-multipoint (PMP), pictured in Figure 4.1. Subscriber stations (SSs), e.g. laptops, handhelds, cars or residences, get access to the network through an association with a base station (BS), pictured as a tower. The association physically takes the form of a wireless link. Logically, the WiMAX/802.16 service, at the link layer, is connection oriented. There are two categories of connections: management connections and transport connections. The former is used for control purposes while the latter is used to carry data traffic.

In contrast to WiFi/802.11, WiMAX/802.16 is designed for long-range wireless access covering several kilometers. WiFi/802.11 is designed for ranges in the order of about a hundred of meters. WiMAX/802.16 is predominantly deployed by cellular or Internet service providers. Typically, WiFi/802.11 networks are used to cover smaller areas ranging from small hotspots in shops to larger areas such as an organization's campus. The WiMAX/802.16 radios primarily operate in the licensed spectrum, although there is support for unlicensed operation. WiFi/802.11 operates exclusively in unlicensed spectrum ranges. The data rates of WiMAX/802.16 and WiFi/802.11 are both in the Mbps range. However, the WiFi/802.11 bandwidth is shared by all devices attached to the access point while the data rate is exclusive for WiMAX/802.16 devices. This results from WiFi/802.11 network access being contention-based. Devices

in the network use carrier sense multiple access with collision avoidance (CSMA/CA) to compete for the channel and transmit. WiMAX/802.16 devices must establish a connection with a BS, competing a single time during network entry. Once connected with a BS, all transmissions are scheduled by the BS. This leads to a much better quality of service (QoS) support in WiMAX/802.16 networks.

WiMAX/802.16 and WiFi/802.11 should be thought of as complementary technologies that can co-exist. WiMAX/802.16 is better suited for metropolitan area networks (MANs). WiFi/802.11 is better suited for local area networks (LANs). One possible scenario is to have WiMAX/802.16 providing the backbone Internet for WiFi/802.11 networks.

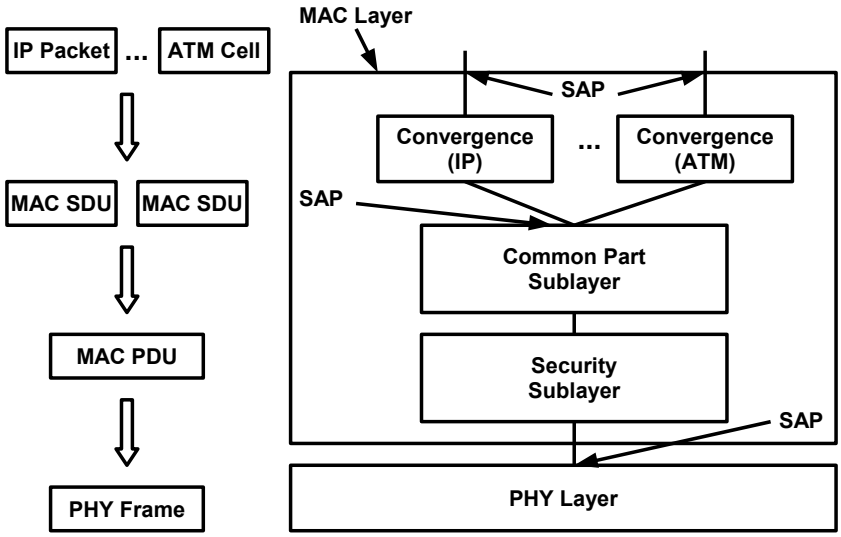


Figure 4.2: Layered architecture of WiMAX/802.16.

The WiMAX/802.16 network architecture is structured into two main layers: the physical (PHY) layer and the medium access control (MAC) layer. The PHY is responsible for the transmission of streams of bits as electromagnetic signals. It is also responsible for the organization of the bit streams into periodic units of fixed lengths called frames. Frames are further divided into smaller units called bursts.

The physical layer is a two-way mapping between MAC protocol data units (PDUs) and physical layer frames received and transmitted through coding and modulation of an RF signal. The clients of the MAC layer pushes data units such as IP packets or ATM cells at identified ports of the WiMAX/802.16 service interface called service access points (SAPs). The MAC layer is subdivided into three sublayers as shown in Figure 4.2. The convergence sublayer adapts data units, accepted through the SAPs, to the requirements of the MAC service data units (SDUs), and vice versa. That might involve, for instance, tasks such as fragmentation and reassembly. There could be several instances of the convergence sublayer according to the kind of adaptation required.

The convergence sublayer also de-multiplexes the MAC SDUs to the connections to which they belong. MAC SDUs are pushed on the common part sublayer. This layer contains the core of the WiMAX/802.16 logic. For instance, it encompasses all aspects of connection management. It creates the protocol data units (PDUs) that are mapped to frames by the PHY layer. The security sublayer addresses authentication, access control and confidentiality issues of the MAC layer traffic.

For the remaining of this chapter, we will first discuss an overview of the WiMAX/ 802.16 MAC layer, including more recent developments such as mobility support and multi-hop relay, in order to give the reader a background into its operation. We will then discuss some previous work done in relation to the scanning and ranging processes.

4.2 Physical Layer

Four choices are available at the PHY layer: WirelessMAN SC, SCa, OFDM or OFDMA. WirelessMAN-SC uses a single radio carrier. It is intended for frequencies in the range of 11 to 66 GHz. It requires line-of-sight (LOS). WirelessMAN-SCa uses a single-carrier at frequencies below 11 GHz and non-line-of-sight (NLOS). Wireless-MAN orthogonal frequency division multiplexing (OFDM) and Orthogonal Frequency Division Multiple Access (OFDMA) are NLOS at frequencies below 11 GHz.

Point-to-multipoint communications consist of a BS and several SSs. The down-link, i.e. the BS to SSs channel, is determined by a frequency and a sector. The

uplink, i.e. the SSs to BS channel, is determined by a frequency. If the downlink frequency and uplink frequencies are the same, then time division duplexing (TDD) is used to control the access to the channel. If the frequencies are different, then frequency division duplexing (FDD) is used. In addition, a time division multiple access (TDMA) technique is used on the uplink.

The flow of bits is structured as a sequence of frames of equal length. There is a downlink subframe and an uplink subframe. In FDD the downlink subframe and uplink subframe are simultaneous, but don't interfere because they are sent on different frequencies. In TDD, the downlink subframe and uplink subframe are consecutive and alternate. Various frame durations are supported between 2 and 20ms in length. In TDD, the portion allocated for the downlink and portion allocated to the uplink may vary. A downlink subframe consists of a preamble, maps and data bursts (see Figure 4.3). The preamble is a sequence of bits used for synchronization purposes. There are two maps. A downlink map announces the start position and transmission characteristics of the following data bursts. An uplink map announces the allocation of the bandwidth to the SSs for their transmission.

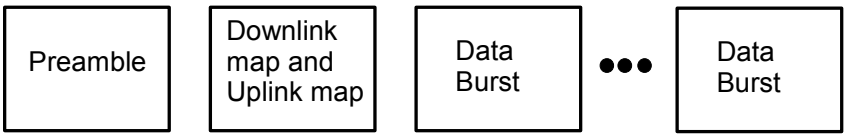


Figure 4.3: A TDD downlink subframe.

In the sequence of data bursts, each burst is transmitted according to a kind of modulation and a kind of forward error correction (FEC). They are sent in increasing degree of decoding difficulty, that is, in increasing level of data rate. Hence, an SS may only decode the bursts up to its reception capabilities and may ignore the bursts it cannot demodulate.

Table 4.1 gives the transmission characteristics of three modulation schemes.

	Data Rate (Mbps)		
Channel Bandwidth (MHz)	QPSK	16-QAM	64-QAM
20	32	64	96
25	40	80	120
30	44.8	89.6	134.4

Table 4.1: WiMAX data rates example.

Quadrature phase-shift keying (QPSK) uses four different symbols and can hence encodes two bits per symbol. 16-quadrature amplitude modulation (16-QAM) uses 16 different symbols and encode four bits per symbol. It doubles the data rate of QPSK. 64-quadrature amplitude modulation (64-QAM) uses 64 different symbols and can encode six bits per symbol. It triples the data rate of QPSK. The degree of decoding difficulty increases with the data rate, because the received signal needs to be stronger. In a downlink subframe, data bursts encoded with QPSK would be transmitted first, followed by the 16-QAM bursts and finally the 64-QAM data bursts.

A FDD downlink subframe contains a TDMA portion assigned to half duplex SSs. The TDMA portion consists of data bursts. Each data burst is assigned to a half duplex SS. A half duplex SS can transmit in a BS assigned TDMA data burst. Also, it needs to receive only when the BS transmits its data burst. A TDMA data burst on the downlink is prefixed by a preamble, transmitted by the BS, to allow the target SS to synchronize. The start position of every TDMA data burst is defined in the downlink map. QPSK is used to send the preambles.

The exact structure of an uplink subframe is determined by the BS. An uplink subframe is made of an arrangement of three kinds of bursts, which may be repeated and appear in any order. The three kinds of bursts are ranging, bandwidth requests and transmission. They are all for SS transmissions destined to the BS. A ranging burst is multiple contention-based access. When an SS initializes its attachment to the network, the SS sends range request messages to the BS with increasing power levels until reception is acknowledged by the BS, which is done by returning a range response message. The requests are sent within a ranging burst. A bandwidth request burst is multiple access contention based, but access must be in response to polling initiated by the BS. An SS may send a request for more bandwidth within such a

burst. A transmission burst is a time interval granted to an SS by the BS.

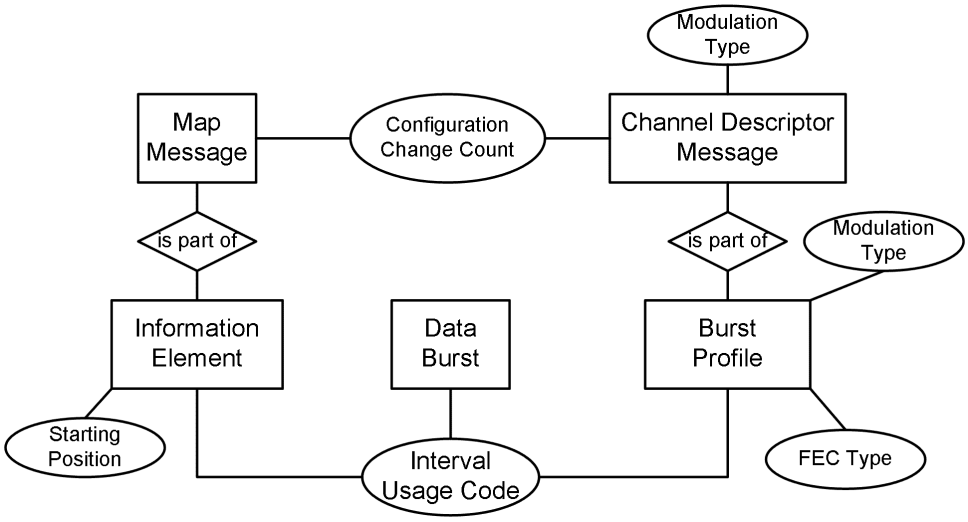


Figure 4.4: Model of PHY and MAC concepts.

The model of information of Figure 4.4 represents the relationships existing among three key entities, that is, data burst, map message and channel descriptor message. Entities are represented by rectangles. Ovals represent their attributes. Lines depict relationships between entities. Data bursts are in downlink subframes and uplink subframes. A data burst is an interval of time during which data is transmitted. A data burst is identified by an interval usage code. A map message is either of type downlink or uplink and prefixes every downlink subframe. It is versioned with a number called a configuration change count. The version number matches the one of the corresponding channel descriptor message. A map message has a BS identifier. It contains a number of information elements. Each information element contains the starting position of a data burst, which is associated with the interval usage code. A channel descriptor message is either of type downlink or uplink. It is sent periodically (maximum of 10 seconds between messages) in a MAC PDU. It is versioned with a number called a configuration change count. The version number matches the one of the corresponding map message. A channel descriptor message contains burst profiles. Each burst profile specifies the transmission characteristics of a data burst, which is associated to an information element with an interval usage code. The specification includes a modulation type and a FEC type.

OFDM is a transmission technique used to share the radio spectrum by several users. Another goal of OFDM is to construct communication channels using, for each of them, several segments distributed over a radio spectrum. This spreading mitigates, on the channel, the undesirable effects of fading or interference that may be stronger on some spectrum segments and weaker on others; it averages them in a sense.

The radio spectrum is divided into N adjacent subchannel elements. Each subchannel element is defined by a radio frequency that is used by a carrier. A channel consists of a number of subchannel elements. Each channel has its specific modulation parameters (rate, coding). The set of subchannel elements is divided into N_G groups. Each group contains N_E subchannel elements, i.e. N is equal to N_G times N_E. Each channel is constructed as the random selection of subchannel elements, each from different groups. This is depicted in Chapter 7, Figure 7.1.

4.3 Medium Access Control Layer

The PHY layer is responsible for sending bits of data from transmitters to receivers. It has no knowledge of the types of data that are transferred between parties. This is the responsibility of the MAC layer. The MAC layer controls access to the PHY layer from the higher layers in the protocol stack. The MAC layer has several other responsibilities including:

- building Protocol Data Units (PDU) to be sent via the PHY layer,
- scheduling transmissions over the PHY layer,
- selecting PHY burst profiles and transmit power levels,
- providing Quality of Service (QoS),
- retransmitting PDUs in case of error, and
- providing support for mobility.

Message Type	Usage
DL basic and UL basic	Short and urgent management messages
DL primary and UL primary	Delay tolerant management messages
DL secondary and UL secondary	IP encapsulated management messages

Table 4.2: MAC layer management connections.

4.3.1 MAC Layer Concepts

Figure 4.2 shows the three components of the MAC layer. The central element is the common part sublayer. It exchanges SDUs with the convergence sublayer. It also constructs PDUs, establishes connections and manages bandwidth. The common part is tightly integrated with the security sublayer. The security sublayer addresses authentication, authorization, key establishment and encryption. The security sublayer exchanges PDUs with the PHY layer.

4.3.2 Management Connections

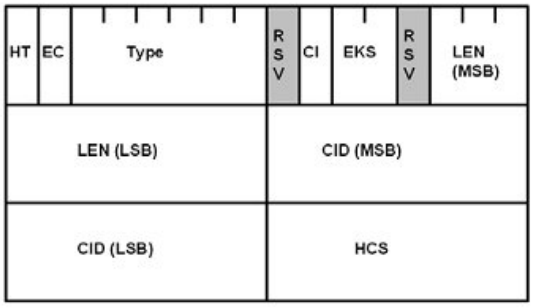
Following the ranging process, two pairs of management connections - one pair member for the uplink and one pair member for the downlink - are established between an SS and a BS. These are called the basic and primary uplink and downlink connections. An optional third pair of management connections, called the secondary management connections, may be created at the end of the network entry. The three pairs of connections handle the different QoS requirements of the management traffic between an SS and a BS. The basic connections are used by an SS and a BS to exchange short, time-urgent management messages, such as requests and responses for changes in burst profiles. The primary connections are used by an SS and a BS to exchange longer, delay-tolerant management messages, such as key management messages. Finally, the secondary management connections are used by an SS and a BS to transfer delay tolerant messages encapsulated in IP packets, such as DHCP, SNMP or TFP messages. Table 4.2 summarizes the management connections.

4.3.3 MAC Protocol Data Units

The common part sublayer is responsible for building PDUs with the SDUs received from the convergence sublayer. Depending on size, a SDU may be fragmented among multiple PDUs or several SDUs may be encapsulated inside a single PDU. There are two types of PDUs: the generic PDU and bandwidth request PDU.

The generic PDU carries user data and MAC layer control messages. The generic PDU contains a header, shown in Figure 4.5, followed by a payload and a cyclic redundancy check (CRC) for error detection purposes. The bandwidth request PDU is used by SSs to request additional uplink bandwidth from a BS. It consists of a bandwidth request header, without payload and CRC.

The bandwidth request header is similar to the generic header shown in Figure 4.5. The header is modified so that the 19 bits, starting at bit 5 and ending at the bit before the CID, holds the bandwidth request (BR) field. The BR field holds the number of bytes an SS is requesting from a BS. The header type (HT) field is set to one. The encryption control (EC) field is set to zero.



Field	Description
HT	Header type(set to 0)
EC	Encryption control (0=not encrypted, 1=encrypted)
Type	Type
CI	CRC Indicator (0=not present, 1=present)
EKS	Encryption key sequence
LEN	Length (including header)
CID	Connection ID for sending PDU
HCS	Header check cequence

Figure 4.5: The generic MAC header.

4.3.4 Bandwidth Request and Allocation

There are several messages sent by a BS in the downlink subframes. These include the downlink and uplink maps (DL-MAP and UL-MAP) and the downlink and uplink

channel descriptors (DCD and UCD). The DL-MAP message defines the start time of every burst in the TDD downlink subframe and its TDMA portion. The UL-MAP message defines the structure of the uplink subframe by giving the offset of every burst. The DCD and UCD describe the PHY layer characteristics for the downlink and uplink.

Requests are sent by the SSs to ask for new uplink allocation from the BS. A request may be in the form of a standalone bandwidth request header or it may be a piggyback request, an optional feature. All bandwidth requests are made in terms of the number of bytes required to contain a MAC header and payload. The request may be transmitted during any allocated uplink slot, but not during the initial ranging interval. Bandwidth requests may be aggregate or incremental. If the BS receives an aggregate bandwidth request, then the number of bytes requested will be the new bandwidth allocation. If the BS receives an incremental bandwidth request, then the number of bytes requested will be added to the existing allocation.

A requested bandwidth is for a SS's individual connection, but the bandwidth granted by the BS is addressed to the SS's basic CID, not to individual CIDs. Since an SS does not necessarily know which request is being granted, it must either assign the bandwidth to its outstanding requests or decide to do a backoff and send additional bandwidth requests if short of bandwidth.

Polling is a mechanism where SSs are allocated bandwidth by the BS in order for them to make bandwidth requests. The allocations are not sent in an explicit message, but are contained as information elements (IEs) inside the UL-MAP. Polling may be done using unicast, multicast or broadcast. The UL-MAP and DL-MAP indicate the uplink and downlink allocations for all the SSs currently communicating with a BS.

4.3.5 Service Flows

Service flows define the parameters of transport connections. They are an important part of the bandwidth allocation process. An SS requests uplink bandwidth on a per connection basis, identifying the service flow, and the BS grants bandwidth as an aggregate of all the per connection requests. Service flows are characterized by a set of QoS parameters. These include latency, jitter and guaranteed throughput

parameters. A service flow has a service flow ID (SFID).

The MAC layer maintains information about provisioned service flows that are used to create connections. This information is stored in the service provisioning tables. From the provisioned service flow table, using the service class index, the QoS parameters for a flow are looked up in the service class table. These parameters include priority, minimum and peak rates, maximum burst, jitter and latency. The classifier rule table provides the rules for mapping IP packets to service flows. Their IP addresses and type of service value determine the service flow, using the service flow index.

4.3.6 Quality of Service

One of the distinguishing features of the MAC layer is the built-in QoS support for five kinds of service flows. An unsolicited grant service (UGS) flow transports real-time traffic consisting of periodic fixed size packets. This is intended for T1/E1 or VoIP traffic. Bandwidth grants are of fixed size. An SS is not required to explicitly request bandwidth. A real-time polling service (rtPS) flow transports real-time traffic consisting of periodic variable size packets. This is suited for MPEG video traffic. To guarantee that the real-time needs are met, a BS provides real-time, periodic unicast request opportunities allowing an SS to specify the size of required bandwidth grants. An extended real-time polling service (ertPS) flow is an enhancement of both UGS and rtPS with unsolicited unicast bandwidth grants (as for UGS), while allowing variable size packets. This intended for voice over IP services with silence suppression. A BS also allows an SS to request additional bandwidth as required. A non-real-time polling service (nrtPS) flow has unicast polls, on a regular basis, to ensure that the SS has have bandwidth request opportunities even when the network is congested. It is intended for applications such as file transfers. A best effort (BE) flow is designed for uplink traffic that has no QoS requirements.

4.4 Network Entry

An SS wanting to establish a connection with a BS follows the network entry procedure depicted in Figure 8. The SS first scans to find a frequency in use by a BS.

It listens to each possible frequency until it hears a frame preamble. The SS synchronizes with the BS by waiting for the DL-MAP and obtains the uplink channel characteristics from the UCD messages that are periodically broadcast by the BS. The SS then gets an Initial Ranging Interval from a DL-MAP message sent by the BS in order to perform initial ranging with the BS. Initial ranging determines the transmit power requirements and timing offsets of the SS in order to synchronize communication with the BS.

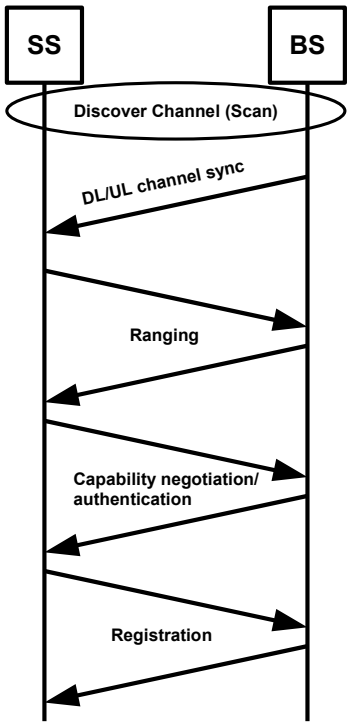


Figure 4.6: Network entry procedure.

4.4.1 Scanning

The goal of scanning is to acquire a downlink signal from a BS. An SS performs scanning by monitoring each possible frequency until a downlink signal is received.

The exact number of frequencies depends on the regulatory provisioned bandwidth. It varies from one country to another. The number of frequencies is also determined by the choice of physical layer specification. Scanning continues periodically to aid SSs in the selection of suitable target BSs for a possible handover in order to maintain network connectivity in mobile WiMAX.

4.4.2 Ranging

The initial ranging process, executed by an SS, determines the transmit power level and timing offset. The timing offset is used to synchronize each SS's transmission to a symbol marking the beginning of a minislot boundary for the SC and SCa PHY layers, or align its transmission with the BS received frame for the OFDM(A) PHY layers. Figure 4.7 shows the initial ranging procedure.

Once the SS has synchronized with the downlink channel and discovered the uplink channel parameters in the UCD message, it uses the UL-MAP message to determine the next initial ranging interval. Next, the SS sends a range request (RNG-REQ) message during the next interval with a CID set to zero. For OFDMA, the request is sent via code division multiple access (CDMA) using codes for initial ranging. The SS adjusts its transmission timings to account for delays, so that it appears to be collocated with the BS. The SS calculates the maximum signal strength P used for initial ranging as follows:

$$P = Max + (EIRP - RSS) \tag{4.1}$$

Max is the maximum equivalent isotropic received power, for the BS receiver. $EIRP$ is the equivalent isotropic radiated power of the BS. Both are obtained from the DCD. RSS is the received signal strength measured at the SS. If the values of Max and $EIRP$ are unknown, then the SS starts sending the RNG-REQ message at a minimum power level, retransmitting and increasing power, during future ranging opportunities after a backoff period, until a BS receives the message. Once a BS has received a RNG-REQ message, it must return a range response (RNG-RSP) message

with the same CID. The RNG-RSP message includes the basic and primary management CIDs assigned to the SS, as well as information on power level adjustment, offset frequency adjustment and timing offset corrections. If the status of the RNG-RSP is success, then the ranging procedure completes. If not, then the SS uses scheduled ranging intervals to complete the ranging process.

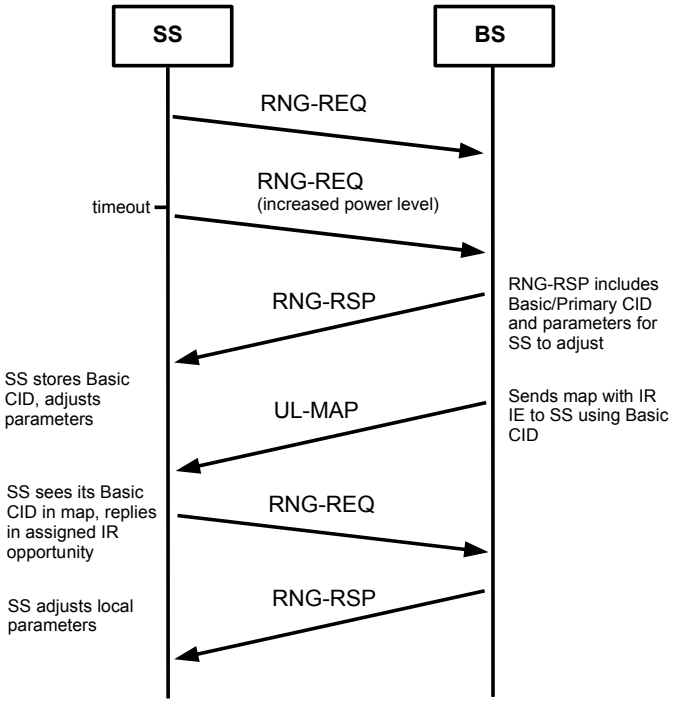


Figure 4.7: Initial ranging process.

4.4.3 Negotiating Basic Capabilities

Negotiating basic capabilities is a necessary step because WiMAX/802.16 is broad, with core functionality that must be supported by all vendors along with optional functionality that different vendors may or may not support. In order for equipment from different vendors to operate together, they must negotiate the set of capabilities

that the BS and SS both support. An example is what PHY burst profiles (e.g. 64 QAM) are supported. The negotiation is a handshake beginning with the SS sending a basic capabilities request (SBC-REQ) message to the BS indicating its capabilities. The BS replies with a basic capabilities response (SBC-RSP) message that includes the capabilities common to both the SS and BS.

4.4.4 Registration

Registration is the process of allowing an SS to enter the network and to receive a secondary management CID. In order to register with the network, the SS sends a registration request (REG-REQ) message to the BS. This message includes information such as uplink CID support, IP version, SS capabilities, convergence sub layer support, ARQ settings and vender specific information. The BS replies with a registration response (REG-RSP) message that includes the secondary management CID. At this point, the SS can establish IP connectivity via DHCP over the secondary management connection.

4.4.5 Establishing a Service Flow

Service flows can be established either from the SS or BS side depending on whether the traffic arriving is on the uplink or downlink. The BS checks to determine if the SS is authorized for service and the QoS requirements can be met with the available resources. It then creates a new SFID with the required class of service and sends this information in a dynamic service addition request (DSA-REQ) message to the SS. If the SS supports the service, then it will respond using a dynamic service addition response (DSA-RSP) message. The SS also reserves the required resources and enables downlink reception. When it receives the response, the BS enables uplink reception. Finally, the BS completes the procedure by sending an acknowledgement (DSA-ACK) message. Both the BS and SS are now ready for reception and transmission can start in both directions.

When an SS wishes to establish a service flow it sends the BS a DSA-REQ message. Upon reception, the BS checks the integrity of the message and responds with a dynamic service received (DSX-RVD) message. The BS then determines if the SS

is authorized for the service and if there are resources available to meet the QoS parameters. The BS then creates a new SFID with the required class of service and sends this in a DSA-RSP response message to the SS. The final step is for the SS to acknowledge the response using a DSA-ACK acknowledgement message. Both the BS and SS are now ready for reception and transmission can start in both directions.

4.5 Mobile WiMAX and Handovers

WiMAX/802.16 supports mobility. Mobility specific issues for the PHY layer, network support for mobility and handovers are discussed hereafter.

4.5.1 Physical Layer

For mobility support purposes, the OFDMA PHY layer has been adapted to use SOFDMA (scalable OFDMA). Here the number of subchannel elements can be chosen to optimize performance according to the channel and the radio conditions. The number of subchannel elements can be 128, 512, 1,024 or 2,048.

4.5.2 Determining Network Topology

To aid mobile stations (MSs) in discovering neighboring BSs, the BSs periodically broadcast information about the network topology using neighbor advertisement (MOB_NBR-ADV) messages. BSs can obtain this information over the backbone network. These advertisements provide the MS with information that they would otherwise obtain by scanning.

BSs assign time for MSs to scan for neighboring BSs. The time that an MS has allocated for scanning is called the scanning interval. A BS may also schedule time to perform interleaving scanning for neighboring BSs. Interleaving scanning is the alternation of sending data and scanning. The MS requests a scanning interval by sending a MOB_SCN-REQ message to the BS including a requested amount of time. The BS responds with a MOB_SCN-RSP message either granting or denying the request. When the MS has been granted a scanning opportunity, it starts scanning for one or more neighboring BS during the time allotted, beginning at the start frame

indicated in the response message.

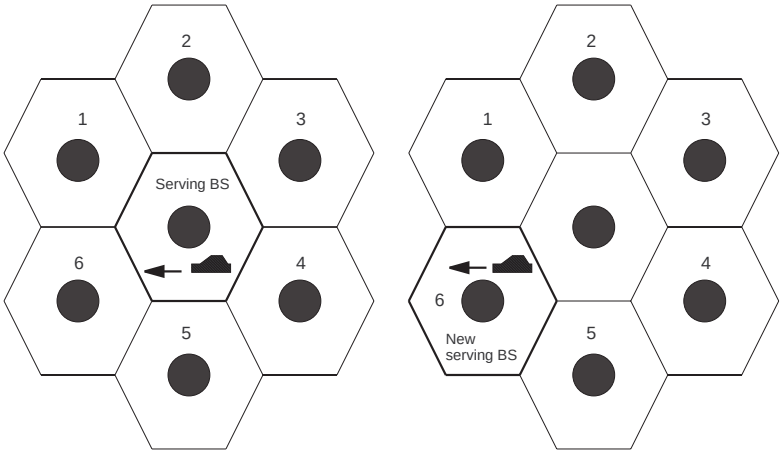
4.5.3 Association Procedure

Association is an optional initial ranging operation that can be done during the scanning interval. This is done between the MS and one of its neighboring BSs. There are three levels of association possible during the scanning interval.

- Level 0: (scan/association without coordination) - The MS performs ranging without assistance from the network. Ranging is contention based. It is successful if the MS receives the success RNG-RSP message from the neighboring BS.
- Level 1: (association with coordination) - The serving BS provides the MS with association parameters (ranging codes and transmission opportunity) for target BSs. This allows the MS to perform collision free ranging and succeeds when it receives the success RNG-RSP message from a neighboring BS.
- Level 2: (network assisted association reporting) - Similar to Level 1, but after the MS has sent its RNG-REQ message it does not need to wait for the RNG-RSP message from a BS. Instead, the RNG-RSP containing information about the PHY layer timing offsets of neighboring BSs is sent to the serving BS over the backbone network. The serving BS may send an aggregation of all this information to the MS in an association report (MOB_ASC-REP) message.

4.5.4 Handovers

As an MS moves throughout the coverage area, maintaining connectivity is done via performing handovers between neighboring BSs. An example of a handover, where an MS must choose one of six neighboring BSs, is shown in Figure 4.8. In this case neighbor six is chosen. Selection of the best handover target can be complex, since the MS must scan for neighboring BSs to find a suitable target based on a number of criteria such as signal strength or error rates. Since a handover is an important function, an MS should perform the scanning and determine a target BS before beginning the handover. Uplink and downlink communication can be temporarily suspended, between an MS and a BS, in order to allow the MS to perform scanning for neighboring BSs. While communication is suspended, the data streams must be buffered on either side.



(a) MS before handover. (b) MS after handover.

Figure 4.8: Mobile station handover from serving BS to neighboring BS 6.

4.5.5 Handover Process

A handover is a process during which an MS changes its connections from one BS to another neighboring BS as shown in Figure 4.8. From the perspective of an MS, the handover process is comprised of a number of steps including cell reselection, handover decision and initiation, synchronization with a target BS, ranging, negotiating capabilities and authentication.

Network assisted handovers are supported in mobile WiMAX. Every BS obtains information about neighboring BSs over the backbone network. The BS periodically sends that information as a mobility neighbor advertisement (MOB_NBR-ADV) message to the MSs. For cell reselection, an MS may use information acquired from the MOB_NBR-ADV messages or may make a request to schedule scanning intervals to scan and perform ranging with a potential target neighboring BS. The handover decision and initiation may originate either at the MS or the serving BS. The notification of intent to handover is sent in either a mobility MS handover request (MOB_MSHO-REQ) message from the MS or a mobility BS handover request (MOB_BSHO-REQ) message from the BS. Scanning and obtaining the downlink and uplink transmission parameters of the neighboring BSs allows an MS to synchronize with the target neighbor BS. If the MS has received a MOB_NBR-ADV message including a potential

target BSs identifier, frequency and channel descriptors, then this procedure can be shortened.

It is possible that a target neighboring BS has received a handover notification from the serving BS over the backbone allowing the target to assign the MS a scheduled ranging opportunity. The MS and target BS must conduct the ranging process in order to determine transmission power level and timing offsets. If the MS includes the serving BS identifier in the RNG-REQ message, then the target BS may use the backbone network to request information about the MS from the serving BS. The target BS can then decide which of the remaining network entry steps may be skipped.

Chapter 5

Strategies for Fast Scanning, Ranging and Handovers in WiMAX/802.16

This chapter describes the strategies we developed in order to make MSs more intelligent during the network entry process. We use the past experiences of an MS to build a list of most likely network parameters to use when trying to find a downlink channel and determining the correct power and timing offsets for communicating with a BS.

5.1 Introduction

In this section we will reintroduce the network entry process and the scanning and ranging procedures as well as discuss why improving the scanning and ranging process is important to improving handover times. Scanning and ranging are performed during network entry and during the handover process. During a handover, the MS must leave its current BS and re-enter the network at a neighboring BS.

5.1.1 Network Entry and Handovers in WiMAX

In order to enter the network, an MS must scan for a frequency in use at the BS. Scanning is an activity conducted by an MS. The goal of scanning is to acquire a downlink signal from a BS. Scanning is done by monitoring each possible frequency until a downlink signal is received. Scanning is performed during the initial network entry procedure and continues periodically to aid the MS in the selection of a suitable target BS for a handover to maintain network connectivity while in motion. It is important that the handover process completes as fast as possible since during the handover data to and from the MS must be buffered. A recent paper found that scanning is the cause of 90% of the delay during handovers in IEEE 802.11r networks [30]

The initial ranging process at an MS involves determining the correct power levels and timing offset for transmitting to the BS. This is to allow MSs to align their transmissions to a symbol marking the beginning of a minislot boundary in the SC and SCa PHYs or aligned with the BS received frame for the OFDMA PHY.

The WiMAX/802.16 specifications provide support for network assisted handovers where the BS currently serving the MS can obtain the information of neighboring BSs over the network. The serving BS periodically sends using broadcast this information as a mobility neighbor advertisement (MOB_NBR-ADV) message to the MS.

Even though there is support for network assisted handovers, the information provided by the base stations has no context. That is, the MS is provided with a list of neighboring BSs and network parameters, but this information does not take into consideration the direction of travel of the MS. Other scenarios where the ability of an MS to make its own decisions on scanning improves performance include the following:

1. Initial Network Entry: here, the MS is not aware of its closest BS and must determine which frequency to use.

2. Handover from one BS to another BS where there is no network handover assistance provided - either it is not available, or the MS is moving between different service providers.

3. Vertical handovers between different network types (i.e. IEEE 802.11 to IEEE 802.16).

5.1.2 Previous Work related to Scanning, Ranging and Handovers

Scanning

The authors in [104] introduce their Adaptive Channel Scanning (ACS) algorithm. ACS is primarily focused on when to perform scanning by estimating the time required for an MS to scan a list of neighboring BSs and then interleaving the scanning and data transmission intervals.

The authors of ACS note that the channel scanning time depends on the synchronization and the association latencies related to the MS association level. They let

l_s and l_a represent the synchronization and association latencies, respectively. The value of l_s will be the time of two frame cycles and the value of l_a is based on the number of frames the MS must wait before sending its range request plus a timeout value. The number of frames, n_f, that an MS must wait is

$$n_f = \left\lceil \frac{M}{n_{cs}} \right\rceil$$

where M is the number of backoff contention slots that the MS must wait and n_{cs} is the number of contention slots per frame. l_a is then $n_f \times t_f + t_{out}$, where t_f is the frame duration and t_{out} the timeout value. The channel scanning time at a BS is $t_s = l_s + l_a$. Once it has determined how long the scanning process will take, ACS uses this information to interleave the scanning operation with ongoing data transmissions.

The authors in [35] propose an optimization of MAC layer management messages for networks that have relay stations. They introduce new management messages for coordination information requests (MOB_SCN-CIR) and association and coordination information responses (MOB_SCN-ACI). For association level 1 and level 2, the BS/RS will transmit the MOB_SCN-CIR message to all recommended stations. This is routed to the serving BS that creates individual messages for each RS and includes the number of recommended stations. If the coordination message is routed trough a recommended station with association level 1 and level 2 it can reply directly to the BS with a MOB_SCN-ACI message and it must reduce the count of recommended stations by one.

Ranging

In their work [90, 91], the authors introduce a BS based algorithm for OFDMA initial ranging that performed well with several users ranging per UL frame. They divided the process into three parts, detection of OFDMA symbols, timing offset estimation and multiuser code detection. They use energy detectors in order to detect OFDMA symbols that contain ranging codes. This allows them to find which symbols to work on and which to drop in order to not waste time on processing empty OFDMA symbols. They then calculate the timing offsets by estimating the linear phase shift for

codes in each OFDMA symbol. Finally, they determine the codes sent by removing the linear phase shift from each transmitted code and cross-correlating the codes against all possible codes.

In [121], the authors propose a new ranging signal detection scheme and an iterative multi-user timing offset estimator. They do not consider collisions of ranging stations. The authors in [122] improve upon the work by [121] by taking into consideration the possible carrier frequency offsets between the received signals and the base station local reference.

Handovers

Other work has been focused more specifically on improving handover performance. The authors in [58] have introduced a new management message to enable the reception of downlink data during the handover process and thus reduce the downlink packet delay. Their proposed algorithm allows an MS to receive downlink data from the handover target BS after synchronizing with the downlink channel, but before synchronizing with the uplink channel. A new MAC management message Fast DL_MAP_IE is used. Until the MS has completed the handover process, it keeps its old CID and the BS uses this CID along with the Fast DL_MAP_IE to indicated downlink data is present for the MS.

The authors in [83] proposed Last Packet Marking (LPM) that requires integrating the MAC layer (L2) handover and network layer (L3) handover and both handovers happen concurrently. Either the MS or BS can initiate a handover. Each BS has a BD-ID to BS-IP mapping table of all neighboring BSs. The currently serving BS then sends a pre-notification message to a target BS for handover, which the target BS can accept or reject. The target BS prepares a buffer for the MS and sends a handover pre-notification response back to the serving BS. The IP address of the MS is updated to associate with the new BS and the previous access router (PAR) is sent a pre-routing update for the MS and forwards data packets towards both the serving and target BS. After this initial pre-handover setup it established, the MS performs a WiMAX/802.16 handover. It then sends a mobile IP routing update message to the gateway to stop the bi-casting of the data from the crossover PAR. The target

BS then forwards any buffered messages to the MS.

Finally, the authors in [78] also propose a cross-layer fast mobile IPv6/802.16e handover scheme (CLHS). They send FMIPv6 handover information with the mobile handover indication (MOB_HO-IND) and RNG-REQ WiMAX/802.16 messages. The fast binding update message is sent with the MOB_HO-IND message as a new FBU_MOB_HO-IND message. The fast neighbor advertisement message is sent with the RNG-REQ message as a new FNA_RNG-REQ message.

Other Related Work

The authors in [24] perform an analysis of several different handover decision algorithms and cite the challenges faced with when deploying video applications in mobile WiMAX neworks and dealing with handovers.

An early work describes the storing of information on the most probable used carrier frequencies in cellular networks on the MS [109]. However, the term most probable is not defined and no mechanism is provided for determining the most probable frequencies.

5.1.3 Results of the Chapter

We present several strategies that an MS uses in order to improve scanning times while searching for a downlink from a BS. The strategies attempt to provide faster network access during the network entry phase as well as reduce delay during the handover procedure. The first set of strategies incorporates the history of successful scanning frequencies in order to guide the MS in choosing frequencies for future scanning operations. A second type of strategy improves upon the use of the mobility neighbor advertisement (MOB NBR-ADV) messages sent to the MS by the serving BS that informs the MS of neighboring BSs. An MS builds a history of handovers between BSs and uses this to determine which BS is the most likely neighbor target BS for a handover. Since a MOB NBR-ADV message gives the MS the list of neighbors and their parameters, knowing which BS is most likely the handover target improves the scanning operation.

Additionally, we present several strategies that an MS uses in order to attempt to

reduce the number of ranging messages sent during the initial ranging process when an MS has no indication of an initial power level with which to begin the ranging process. As with our scanning strategies, two of our ranging strategies are based on the history of previous ranging operations.

Related works described in Section 5.1.2 either try to work around the required scanning by estimating how long it will take and then schedule the scanning operation efficiently, implement mechanisms such as new control messages or cross-layer handover techniques to improve upon the impact of scanning, or require additional BS support. We attempt to reduce the time required for the scanning and ranging operations. Our strategies require no additional network support and only limited memory and computational resources of the MS.

We evaluated our strategies by mapping *real-world* mobility trace data for a set of MSs obtained via the Automatic Packet/Position Reporting System (APRS) project by Bruninga to perform MS simulations within our coverage area.

5.2 Model Definition

We assume an area of WiMAX coverage of size A in which an MS is located. All locations are equally probable. We have a set of n frequencies $f_1, ..., f_n$. Each frequency covers a portion of the area A. The size of the area covered by frequency f_i is denoted as $A(f_i)$. We have a frequency reuse cluster size c. The number of frequencies per BS, f_{BS}, is

$$f_{BS} = \frac{n}{c}$$

The area covered by a frequency f_i can be determined by

$$A(f_i) = \frac{A}{\frac{n}{f_{BS}}}$$

We assume an MS obtains a downlink channel on one of the frequencies assigned to a BS with equal probability. Assuming that there is at least one channel available at BS_j, the MS finds that f_i is available with probability

$$p(f_{i,BS_j}) = \frac{1}{c} \times \frac{1}{f_{BS}}$$

The probability of finding the downlink signal at frequency f_i is

$$p(f_i) = \frac{A(f_i)}{A} \times p(f_{i,BS_j})$$

In Figure 6.5 we see the area of WiMAX coverage and the frequency reuse scheme. Here we let $c = 4$.

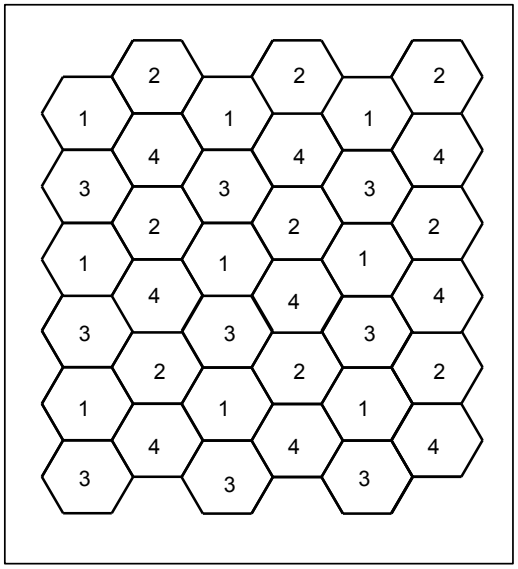


Figure 5.1: WiMAX coverage, frequency use, four cell clusters.

We assume that the time required to scan a frequency corresponds to the duration T of a physical layer scan. If the set of n frequencies are scanned in the order $f_1, f_2, ..., f_n$, then the average time it takes for scanning is expressed as follows

$$E(f_1, ..., f_n) = p(f_1)T + \sum_{i=2}^{n} ip(\neg(f_1, ..., f_{i-1}), f_i)T$$

Where the expected time to acquire the downlink signal during scanning is equal to the probability that the frequency f_1 was found times the time, T, required for scanning a channel plus the sum of for 2 to n that the frequency was found at the i^{th} frequency and not in the previous frequencies. The goal is to find an ordering of the n frequencies such that

$$E(f_1, ..., f_n)$$

is minimum.

5.3 Scanning Strategies

In the following section, we describe the WiMAX/802.16 default scanning strategy as well as two new strategies. The first new strategy, deals with reducing the number of frequencies checked during each scanning operation. The MS maintains information on the probabilities of the frequencies being used in order to make a choice. This strategy is further divided into the *Most Recently Used* and *Most Frequently Used* approaches. The second new strategy deals with the MS maintaining information on the history of performing handovers between BSs that were encountered under the assumption that patterns observed in handovers will be repeated.

5.3.1 Default WiMAX Strategy

The IEEE 802.16 standard [8] specifies that an MS must keep a nonvolatile storage where it stores the last set of operational parameters. When an MS wishes to acquire a downlink channel, it uses this stored information. Whenever an MS fails to obtain the downlink channel, it then continuously scans the possible channels of the downlink frequencies until a time it finds a downlink signal.

5.3.2 Frequency Strategies

When an MS is turned on for the first time, all frequencies are equally likely since the MS has no history. That is to say if there are n frequencies, that the frequencies f_1, f_2, ... , f_n have an initial probability distribution $p(f_1), p(f_2), \ldots, p(f_n) = \frac{1}{n}$. This can be defined as a random variable X assuming the values $f_1, f_2, \ldots f_n$ such that $Pr[X = f_i] := p(f_i)$. An MS only moves along a limited geometric planar region defined by the cells it traverses. This in turn gives rise to a random variable X_{MS} induced by the random variable X and having probability distribution $Pr[X_{MS} = f_i]$. It is clear that this probability distribution depends on the mobility characteristics (the path) of the MS as well as the probability distribution of the frequencies in the

subregion traversed by the MS during its movements. Given this random variable, we are interested in optimizing how the MS selects its frequencies.

From the initial setup, since an MS has no previous history, it must simply start scanning frequencies in increasing order. As the MS performs a number of successful scanning operations, it can determine an *order of frequencies* from its observations. This history can be used to make scanning more efficient.

In the following sections we discuss two new strategies that utilize the MSs history of scanning operations to determine the frequency ordering.

Most Recently Used Strategy (MRU)

In this first strategy, the MS must keep a certain number of the possible frequencies stored in memory. This information is independent of BSs or network topology. Initially, all frequencies are equally likely and are chosen in increasing order (lowest frequency to highest frequency). For each successful scanning operation, the frequency discovered, f_{MRU_1}, is given the highest priority, that is, it is moved to the front of the list of frequencies. As the MS builds up a history of scanning operations they will be ordered as follows

$$f_{MRU_1}, f_{MRU_2}, f_{MRU_3}, \ldots, f_{MRU_k}$$

where f_{MRU_1} is the *most recently used* frequency, f_{MRU_2} is the second most recently used, and so on. The value of k is the number of frequencies stored in memory. There is a training period of time required to establish this order. In future scanning operations, the MS scans frequencies in order from the *most recently used* (MRU) to the least recently used. If the MS does not find the frequency from the MRU list, then it must scan the remaining frequencies. The MRU strategy was first used in relation to paged memory caches [108].

Most Frequently Used Strategy (MFU)

In this strategy, a frequency of occurrence distribution over the frequencies is built based on the history of the number of times each frequency is used. The frequency that is used the most has the highest priority. As the MS builds up a history of scanning operations, the frequency of occurrence, F, distribution is as follows

$$F(f_{MFU_1}) \geq F(f_{MFU_2}) \geq F(f_{MFU_3}) \geq \cdots \geq F(f_{MFU_k})$$

where f_{MFU_1} is the most frequently used frequency, f_{MFU_2} is the second most frequently used and so on. The value of k is the number of frequencies stored in memory. The MS keeps an ordered list starting with the most frequently used frequency. In future scanning operations, the MS scans the frequencies from the most frequently used to the least frequently used frequencies. The MFU strategy was first introduced for use in multi-level memory caching [97].

5.4 Ranging Strategies

In this section, we describe several ranging strategies including ones that can be used in order to reduce the number of ranging messages that an MS must send in order to coordinate the transmission power level and timing offsets with the BS.

Our model defines a set of transmission power levels that are available to the MS. Let the first transmission power level (the lowest power level) be the power required in order to be able to communicate with a BS up to a distance of one unit, the second power level covers a distance of two units, the third covers three units and so on. Based upon the assumption that the energy required to transmit is proportional to the distance squared, if the energy required for level one is one unit, then the energy required for level two is four units, level three is nine units and so on. We can now measure the amount of energy required for the ranging process. The goal of our strategies is to reduce the number of ranging messages sent on average by the MS. However, we must also examine how much energy is expended during an average ranging operation.

Once a BS receives a ranging message from an MS, it notifies the MS in a range response message of the power adjustments required. We consider a ranging operation to be successful when the MS transmits with enough power for the BS to initiate a response.

Default Strategy

While the WiMAX/802.16 standard does not define the ranging strategy to be used by the MS, the simplest strategy is for the MS to start transmitting the ranging messages at the lowest power level, and incrementing through each increasing power level until it receives a range response message from the BS.

Exponential Increment Strategy

This next strategy calls for the MS to start transmitting the ranging messages at the lowest power level, i, but then instead of simply incrementing to the next power level, it moves to the i^2 power level. That is to say, the first attempt is at power level one, the second at power level two, the third attempt at power level four, and so on until the MS receives a range response from the BS. This strategy has the benefit of being a memoryless strategy since it requires no previous knowledge of ranging operations. This strategy also quickly increases the transmit power compared to the default strategy.

Most Recently Used Strategy

In this strategy, the MS keeps a list of a certain number of the possible transmission power levels in memory. This information is independent of BSs or network topology. Initially, all power levels are equally likely and are chosen in increasing order (lowest power level to highest power level). For each ranging operation, the power level used, P_{MRU_1}, is given the highest priority, that is, it is moved to the front of the list of power

levels. As the MS builds up a history of ranging operations, they will be ordered as follows

$$P_{MRU_1}, P_{MRU_2}, P_{MRU_3}, \ldots, P_{MRU_k}$$

where P_{MRU_1} is the *most recently used* power level, P_{MRU_2} is the second most recently used, and so on. The value of k is the number of transmission power levels stored in memory. There is a training period of time required to establish this order. In future ranging operations, the MS sends repeated ranging messages with power levels in the order from the *most recently used* (MRU) to the least recently used until it receives a range response from the BS. If the next power level stored entry in the list is lower than the previous attempt, then it is skipped and the next greater power level in the list is used. If the MS does not receive a range response from the BS, then it sends ranging messages starting with an increment above the highest power level attempted thus far.

Most Frequently Used Strategy

In this strategy, a frequency of occurrence distribution over the transmission power levels used is built based on the history of the number of times each power level is used. The power level that is used the most has the highest priority. As the MS builds up a history of ranging operations, the frequency of occurrence, F, distribution is as follows

$$F(P_{MFU_1}) \geq F(P_{MFU_2}) \geq F(P_{MFU_3}) \geq \cdots \geq F(P_{MFU_k})$$

where P_{MFU_1} is the most frequently used power level, P_{MFU_2} is the second most power level used and so on. The value of k is the number of power levels stored in memory. The MS keeps an ordered list starting with the most frequently used power level. In future ranging operations, the MS sends repeated ranging messages with power levels from the most frequently used to the least frequently used until it receives a range response from the BS. If the next power level stored in the list is lower than the previous attempt it is skipped and the next higher power level in the list is used. If the MS does not receive a range response from the BS, then it sends ranging messages

starting with an increment above the highest power level attempted thus far.

5.5 Previous Handover Strategy

In this section, we describe a new strategy that utilizes the history of the MS handovers along with the information made available from the currently serving BS in the MOB_NBR-ADV message to improve the choice of a handover target neighbor BS to begin scanning.

We assume the MS is provided with the list of all neighboring BSs and the frequency advertised for each neighbor in the MOB_NBR-ADV message from the serving BS. Even though this information is provided to the MS, the MS must "guess" at which of the neighboring BSs it should attempt to perform a handover. The MS must perform a scanning operation to obtain the operational parameters and determine which of the neighboring BSs are suitable for a handover. The IEEE 802.16 standard does not define this operation. It is left to the MS to decide on which neighbor to attempt to communicate. an MS with no other information than the MOB_NBR-ADV must simply attempt to scan for the neighbors in the order given.

That is to say, that the d given neighbors for a given BS are n_1, n_2, ... , n_d and that they are equally likely to be chosen as a handover target. This gives them an initial probability distribution $p(n_1), p(n_2), \ldots, p(n_d) = \frac{1}{d}$. This can be defined as a random variable Y assuming the values $n_1, n_2, \ldots, n_d$ such that $Pr[Y = n_j] := p(n_j)$. We want to capture the probability that the MS currently served by BS_i will perform a handover to BS_is neighbor j. Similar to the frequency strategies, this probability distribution depends on the mobility characteristics (the path) of the MS. Given this random variable, we are interested in optimizing how the MS selects its target handover BS. $p(HO_{i,j})$ is the probability that a handover will occur from BS_i to it's neighbor n_j.

The MS keeps for each BS_i it has visited, a list of most probable handover target of the neighboring BSs. For example, a commuter driving along a highway is highly likely to repeatedly perform the same set of handovers along the highway, from BS_1 to BS_2 to BS_3 and so on. This pattern can be captured and stored in the MS. This can be done by maintaining a frequency of previous handovers that have occurred

from each BS_i. The frequency of occurrence, F, distribution is as follows

$$F(HO_{i,i_1}) \geq F(HO_{i,i_2}) \geq F(HO_{i,i_3}) \geq \cdots \geq F(HO_{i,i_d})$$

where we have non-increasing probabilities and $p(HO_{i,i_1})$ is the probability of a handover from BS_i to one of its neighbors n_{i1} which will be considered first when the MS performs scanning in preparation for a handover.

5.6 A Note on MRU and MFU Strategies

Our *MRU* and *MFU* strategies have a basis from various works done on computer memory cache systems [97, 108].

When a computer is executing code inside a loop it will be going through the exact same set of instructions repeatedly. Therefore, by using *MRU* or *MFU* cache replacement strategies it can improve performance with repeated cache hits. The frequencies used by an MS represent instructions used in a memory cache while the MSs in our model would represent the processes within a computer system. Our model differs since if an MS follows the same path twice, it is not guaranteed to connect to the set of BSs with the same set of frequencies. This is dependent on the number of frequencies allocated to each BS.

Additionally, in memory cache systems, typically there is only a single cache of one or more levels. As such, the cache is used to improve performance in the execution of instructions from *all* processes running on the system. In our initial model, we are implementing our strategies with the goal of building mobility profiles for each individual MS. This will ultimately result in an MS keeping multiple lists of *most likely* used frequencies depending on different conditions such as time-of-day, or current location and heading.

Finally, our strategies involve maintaining an ordered list of most likely frequencies that are traversed from the beginning to the end while searching for BS frequencies. Typically in a memory cache, the systems can directly load the instruction stored in cache and must not traverse the cache looking for it.

The work in this chapter shows that by keeping a single list of most likely frequencies used we can improve our scanning and ranging performance. In the following chapter, we expand these ideas by building multiple lists of most likely used frequencies based upon more complex and fine grained mobility profiles.

5.7 Performance Evaluation

In order to evaluate our strategies, we implemented a simulation to compare the WiMAX/802.16 default scanning strategy with the *MRU* and *MFU* strategies. We then implemented a second simulation to test the *previous handover strategy*. Finally, a third simulation was implemented to compare our ranging strategies. In the following section, we first describe important parameters concerning our simulation: (1) Network topology, (2) Metrics and measurements being used, (3) the Mobile Station Mobility Data used, and (4) Mobile Station scanning and ranging times. Later we discuss the results of the simulation.

5.7.1 Simulation Setup

The simulations are setup as summarized in Table 5.1. A ten kilometres by ten kilometres area is defined and covered by WiMAX/802.16 base stations with a one kilometre range. The base stations are positioned based upon a cellular networking model, see [25], where each BS covers a hexagonal shaped region as shown in Figure 6.7. A total of 44 BSs are required to cover the entire area.

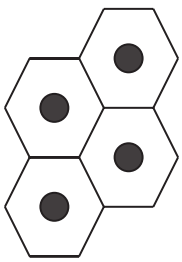


Figure 5.2: Four cell cluster size for frequency reuse.

Network Parameters	
Coverage Area	10 km x 10 km
No. Base Stations	44
Cell Radius	1 km
Cell Cluster Size	4
No. Frequencies	20
No. Channels	80
No. Channels per Cell	20
No. Frequencies per cell	5

Table 5.1: Network simulation parameters.

The WiMAX/802.16 simulation parameters are chosen based on 100 MHz of bandwidth available for both uplink and downlink channels. Every channel is given 1.25 MHz of bandwidth for a total of 80 channels. A four cell cluster reuse clustering scheme, shown in Figure 5.2, is implemented on the underlying 44 BSs. This setup has 20 channels available for each cell for the 80 channels available in total. We use 20 different frequencies, assigning five to each BS, each of which has four channels for MS connections. Two MS scanning times are implemented in the simulation scenario.

In our model, the number of frequencies stored in memory is equal to the number of frequencies available in the coverage area.

For our ranging simulation model, we define the MS power levels relative to our area of coverage. Power level one allows communication between an MS and a BS up to 100 metres apart, power level two allows communication up to 200 metres, and so on.

Metrics and Measurement

The purpose of this work is to improve upon the time it takes for an MS to find a downlink channel and perform initial ranging. For the scanning process, we aim to reduce the number of frequencies an MS is required to check while scanning for a downlink connection to a BS. We measure the proportion of time that an MS can find the downlink with a certain number of frequencies checked. For our model, this is between one and 20 frequencies checked per scan. Ideally, we would like to have it so that 100% of the time an MS can successfully scan for a neighboring BS with

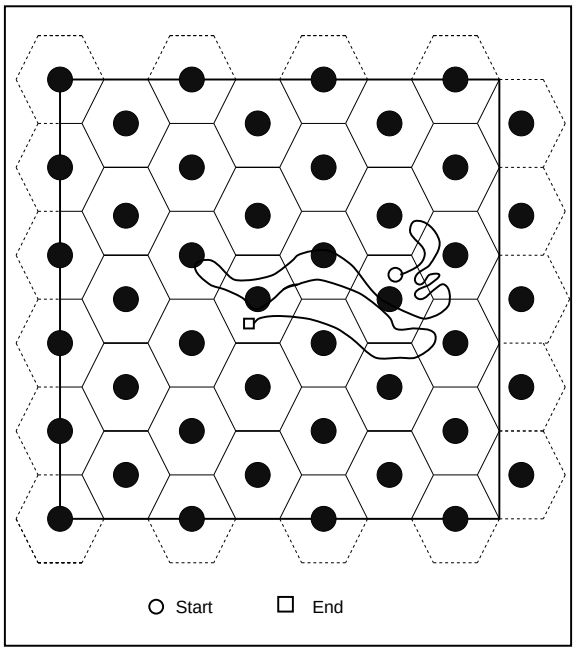


Figure 5.3: Simulation area with sample path of an MS.

only checking a single frequency, but the actual performance is dependent on the individual MSs mobility. From the observations of our simulations, we determine the expected number of frequencies required to be checked for a certain proportion of scanning operations. For example, $x\%$ of the time, the MS can find a downlink channel with only scanning a single frequency. Results were obtained with a 95% level of confidence. Similarly, to measure an improvement in the ranging process, we aim to reduce the number of ranging messages that an MS sends to the BS when performing the ranging process. From our simulations, we determine the proportion of time that the ranging process completes with one, two, up to n ranging messages. Furthermore, we determine the average number of ranging messages sent and average energy use per ranging attempt for each of the ranging strategies.

MS Mobility Data

We used real-world mobility data collected by the APRS project [50]. The data included the mobility traces of 30 different MSs. The actual real-world coordinates of the data for each MS were mapped to the simulation coverage area. As an example, a

partial depiction of an approximated path, in terms of handovers, of an MS is shown in Figure 5.3.

MS Scanning and Ranging Times

Here we describe two different scanning and ranging times of an MS. The first represents how the operations are performed as an MS is in motion and performs handovers in order to maintain a network connection. The second is more representative of the cases where an MS is just being turned on and must find a downlink channel and range to a BS in order to enter the network.

1. *Along the MS path*: In this scenario, we chose sampled locations that were *relatively* close in time to the *previous* scanning and ranging attempt. This would closely follow the path taken by the MS and would cover scenarios where the MS is trying to re-acquire a channel, or is performing a handoff between cells. We expected that the *MRU* strategy would perform better in this scenario. Observe in Figure 5.3 that an MS may frequently move back and forth between the same series of two or more cells. It should be noted that the IEEE 802.16 default strategy is equivalent to the MRU strategy with a list of only one frequency.

2. *Random Locations*: In this scenario, we implement a series of randomly chosen sampled locations of each MS and perform the scanning and ranging operations. This can be thought of to represent when a mobile station is turned on. It was expected that the *MFU* strategy would achieve better performance with this scenario since a user can be expected to perform this kind of operation in certain places (i.e. home, office, etc.)

We performed simulations for the IEEE 802.16 default, *MRU*, and *MFU* scanning strategies as well as the default, *MRU*, *MFU* and *Exponential* ranging strategies in order to compare their results.

Previous Handover Strategy

In the previous handover simulation setup, the same area of coverage and mobility trace data were used. However, now the MS built a history of the previous BS

handover pairs, and store for each BS, the list of neighboring BS in a *MFU* order. That is, for each BS, the list of neighboring BSs is maintained in the order of most often chosen as the handover target to the least often chosen as the handover target. When an MS receives a MOB_NBR-ADV message, it chooses the neighbor that is most frequently used first. From our simulation model, an MS has at most six neighboring BSs from which they can choose for a handover. We performed simulations for the IEEE 802.16 default and the *MFU* handover strategy in order to compare their results.

5.7.2 Simulation Results

The following sections show the simulation results of our frequency, ranging and handover strategies. Additionally, the complete set of frequency scanning tables, including 95% confidence intervals can be found at the end of the chapter.

Scanning Strategies

Figure 5.4 gives the results of the scanning *along the path* and *random locations* scenarios showing the distribution of the proportion of the time that an MS is required to scan a certain number of frequencies before finding the downlink signal from a BS. For the *along the path* scenario, the results show that the *MRU* strategy performs the best of the three strategies.

Table 5.2 summarizes the results and can be read as follows: for our network parameters, the WiMAX/802.16 default scanning strategy finds the downlink 51.9% of the time by checking 10 frequencies. The MRU strategy finds the downlink 86.9% of the time with the same number of frequencies checked. In order to meet the MRU number, the default strategy requires checking 18 frequencies.

For the *random locations* scenario, we observe that the *MFU* strategy performs best. The results are further summarized in Table 5.2 (the values in parentheses) where we see that the MFU strategy successfully finds a BS frequency 79.5% of the time after checking 10 frequencies versus only 45.6% of the time for the WiMAX/802.16 default strategy. In order to meet the 79% of the MFU strategy, the default strategy is required to check 17 frequencies on average.

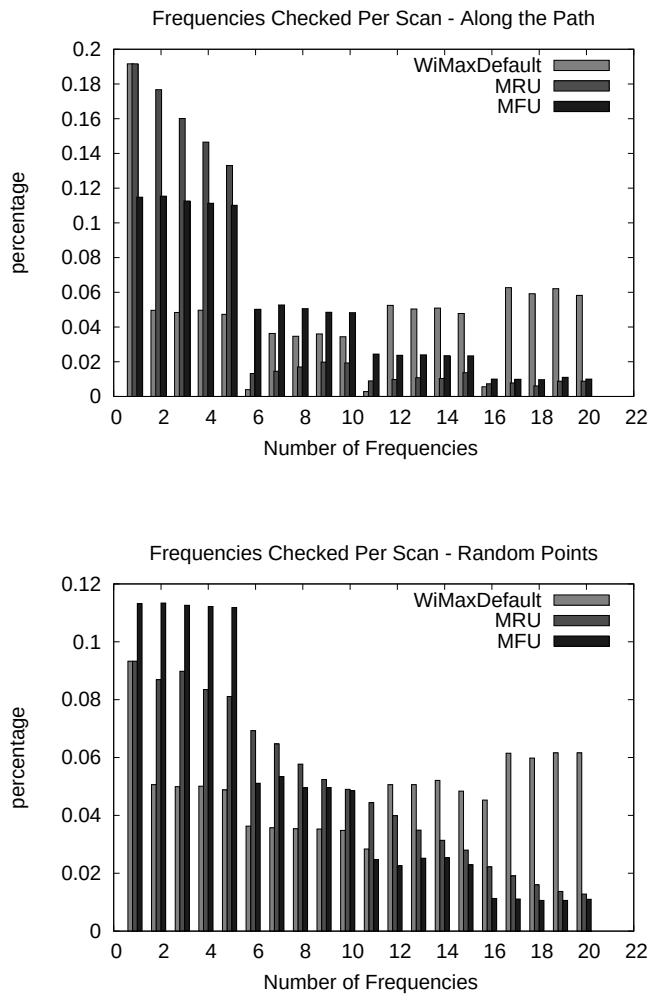


Figure 5.4: Distribution of number of frequencies checked per scan.

Scanning Results		
	5 Freq	**10 Freq**
Default	39.9% (30.3%)	51.9% (45.6%)
MRU	78.0% (40.0%)	86.9% (70.7%)
MFU	53.2% (52.9%)	79.5% (79.5%)

Table 5.2: Proportion of time scanning can be completed with checking 5, 10 frequencies - along the path (random locations).

Ranging Strategies

The following are the results of our ranging simulations. The graph on top in Figure 5.5 shows the results of the *along the path* scenario and shows the distribution of the proportion of the time that an MS is required to send from one to eleven ranging messages (eleven is the maximum due to our model of BS coverage). The results show that the *MRU* strategy performs the best of the strategies. Similarly, the graph on the bottom in Figure 5.5 shows the results for the *random locations* scenario. Here we observed that the *MFU* strategy performs best.

The results are further summarized in Table 5.3 and Table 5.3 where we see the average number of ranging messages and average energy required per ranging attempt. The default strategy required the most average number of ranging messages and average energy consumption, while both the *MRU* and *MFU* strategies performed significantly better in terms of both metrics. The *MRU* strategy performed better for the *along the path* scenario in terms of the average number of ranging messages sent, while the *MFU* strategy performed better in the *random locations* scenario. The *MRU* strategy performed best overall in terms of the average energy expended per ranging attempt. Finally, we observe that the *exponential increment* strategy only performed moderately better than the default strategy. It might be noted that with an increased BS coverage range (currently at one kilometer), the difference in performance between the default and *exponential increment* strategy may be more pronounced.

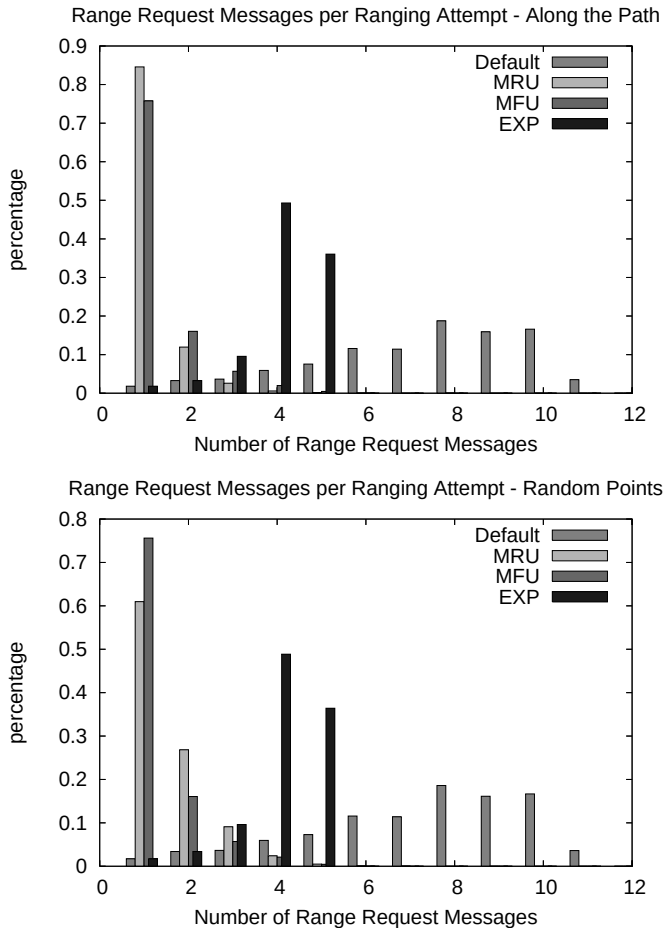


Figure 5.5: Distribution of range request messages per ranging attempt.

Previous Handover Strategy

In this simulation scenario, the WiMAX/802.16 default was to simply scan the neighbors in the order given by the MOB_NBR-ADV messages. By using information on the historical pattern of handovers from each BS, we obtain significant improvement in choice of neighbor BS frequency to scan. The results of the simulation are shown in Figure 5.6. We see that with the default strategy, slightly more than 20% of the time the MS picked the correct neighbor on the first attempt while with the MFU strategy this increases to over 70%. While the WiMAX/802.16 default was more evenly distributed, the MFU strategy successfully finds the correct handover

Ranging Results - Along the Path		
	Avg. Ranging Msgs	**Avg. Energy Expended (units)**
Default	7.28 ± 0.32	198.52 ± 18.81
MRU	1.20 ± 0.04	57.76 ± 4.42
MFU	1.35 ± 0.13	74.77 ± 7.24
EXP	4.14 ± 0.11	167.03 ± 17.15

Table 5.3: Average number of ranging messages and energy expended per ranging attempt - along the path, 95% confidence.

Ranging Results - Random Points		
	Avg. Ranging Msgs	**Avg. Energy Expended (units)**
Default	7.29 ± 0.33	199.30 ± 18.96
MRU	1.55 ± 0.04	53.21 ± 4.16
MFU	1.36 ± 0.13	74.98 ± 7.27
EXP	4.15 ± 0.11	167.87 ± 17.23

Table 5.4: Average number of ranging messages and energy expended per ranging attempt - random points, 95% confidence.

target with only two attempts over 90% of the time.

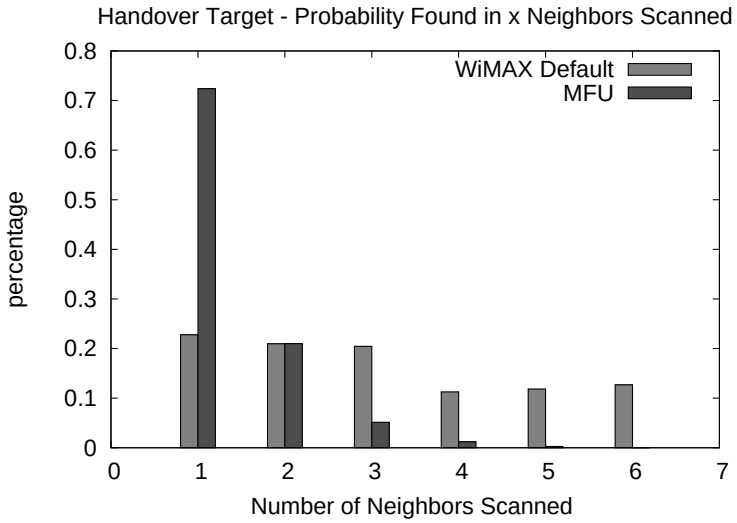


Figure 5.6: Previous handover strategy - BS neighbor scans per handover.

Percentage Scanning Checks Number of BSs		
No. Neighbors Scanned	WiMAX	MFU
1	22.78% ± 6.07%	72.38% ± 6.74%
2	20.96% ± 5.09%	20.98% ± 3.42%
3	20.45% ± 3.94%	5.15% ± 1.28%
4	11.27% ± 2.95%	1.20% ± 0.50%
5	11.84% ± 3.34%	0.26% ± 0.21%
6	12.69% ± 3.33%	0.01% ± 0.03%

Table 5.5: Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs, 95% confidence

5.8 Conclusion and Future Work

We have examined the scanning and ranging operations for a WiMAX/802.16 MS. We have introduced two new strategies to aid in reducing the number of frequencies to check while scanning to find a downlink from a BS and provided some experimental results of a simulation based on real-world mobility traces. In the two scenarios we tested, *along the path* vs. *random locations*, we have shown that the *MRU* (*MFU* respectively) performed best when compared against the default WiMAX/802.16 scanning strategy. The proportion of time that scanning was completed within checking ten frequencies has been increased from 51.9% (45.6%) for the WiMAX default to 86.9% (70.7%) for the *MRU* strategy and 79.5% (79.5%) for the *MFU* strategy for along the path (random points) scenarios.

We introduced an additional strategy to enhance the neighbor advertisement messages sent by the BS to assist the MS with handovers. In this strategy, the MS builds a history of previous handovers and uses this information to select which of the advertised neighboring BSs is most likely the correct one for a handover. We have shown that the *previous handover* strategy improves performance significantly over that of the WiMAX/802.16 default strategy. On average, the *previous handover* strategy chooses the correct target BS on its first attempt 72.38% of the time as compared to only 22.78% of the time for the WiMAX/802.16 default.

Finally, we have introduced a series of initial ranging strategies to aid in reducing the number of ranging messages sent during the ranging process. In the two scenarios

we tested, *along the path* vs. *random locations*, we have shown that the *MRU* (*MFU* respectively) performed best when compared against the default and exponential increment ranging strategies. The average number of ranging messages required has been reduced from 7.28 (7.29) for the WiMAX default to 1.20 (1.55) for the *MRU* strategy and 1.35 (1.36) for the *MFU* strategy for along the path (random points). Additionally, with by using the *MRU* strategy for ranging, the power expended has been reduced by between 70.9% and 73.3% over the default WiMAX strategy.

Our work is complementary to the earlier research presented such as that done by [58, 78, 83, 90, 104]. Our strategies are MS based and require no additional network support. This initial work on scanning, ranging and handovers was the basis for our work on developing the TOD and LPT mobility profiles in Chapter 6. This work has been published in [40] and [41].

5.9 Overall Scanning Results

Scanning Results - Along the Path - APRS Mobility Data			
No. Freq Checked	WiMAX	MFU	MRU
1	19.16% ± 1.20%	11.48% ± 1.15%	19.16% ± 1.20%
2	4.96% ± 1.25%	11.54% ± 1.16%	17.67% ± 1.30%
3	4.84% ± 1.25%	11.26% ± 1.24%	16.01% ± 1.32%
4	4.97% ± 1.27%	11.13% ± 1.23%	14.65% ± 1.28%
5	4.73% ± 1.22%	11.01% ± 1.17%	13.30% ± 1.34%
6	0.39% ± 0.13%	5.02% ± 0.76%	1.32% ± 0.39%
7	3.63% ± 1.23%	5.27% ± 0.85%	1.45% ± 0.32%
8	3.47% ± 1.21%	5.06% ± 0.70%	1.70% ± 0.35%
9	3.60% ± 1.27%	4.85% ± 0.66%	1.98% ± 0.40%
10	3.44% ± 1.22%	4.83% ± 0.72%	1.93% ± 0.36%
11	0.28% ± 0.13%	2.44% ± 0.45%	0.90% ± 0.25%
12	5.25% ± 1.34%	2.37% ± 0.43%	0.98% ± 0.27%
13	5.04% ± 1.33%	2.40% ± 0.46%	1.08% ± 0.32%
14	5.09% ± 1.35%	2.35% ± 0.43%	1.04% ± 0.29%
15	4.78% ± 1.27%	2.34% ± 0.42%	1.37% ± 0.44%
16	0.56% ± 0.21%	1.00% ± 0.27%	0.73% ± 0.37%
17	6.27% ± 1.39%	0.99% ± 0.27%	0.78% ± 0.41%
18	5.92% ± 1.33%	0.97% ± 0.28%	0.61% ± 0.20%
19	6.21% ± 1.40%	1.10% ± 0.34%	0.88% ± 0.31%
20	5.82% ± 1.28%	1.00% ± 0.28%	0.88% ± 0.23%

Table 5.6: Proportion of time scanning can be completed checking between 1 and 20 frequencies - along the path, 95% confidence

Scanning Results - Random Points - APRS Mobility Data			
No. Freq Checked	WiMAX	MFU	MRU
1	9.33% ± 1.08%	11.32% ± 1.14%	9.33% ± 1.08%
2	5.06% ± 1.29%	11.34% ± 1.16%	8.69% ± 1.09%
3	4.99% ± 1.23%	11.26% ± 1.21%	8.98% ± 1.02%
4	5.01% ± 1.25%	11.22% ± 1.20%	8.35% ± 1.02%
5	4.88% ± 1.24%	11.18% ± 1.21%	8.11% ± 0.96%
6	3.63% ± 0.65%	5.11% ± 0.75%	6.93% ± 0.68%
7	3.57% ± 1.27%	5.34% ± 0.90%	6.47% ± 0.60%
8	3.54% ± 1.22%	4.95% ± 0.72%	5.77% ± 0.55%
9	3.53% ± 1.21%	4.96% ± 0.69%	5.24% ± 0.57%
10	3.48% ± 1.22%	4.85% ± 0.69%	4.90% ± 0.56%
11	2.84% ± 0.60%	2.47% ± 0.48%	4.44% ± 0.53%
12	5.06% ± 1.32%	2.26% ± 0.40%	3.99% ± 0.53%
13	5.06% ± 1.30%	2.52% ± 0.51%	3.49% ± 0.48%
14	5.21% ± 1.34%	2.53% ± 0.54%	3.14% ± 0.48%
15	4.84% ± 1.30%	2.29% ± 0.42%	2.80% ± 0.44%
16	4.53% ± 0.84%	1.13% ± 0.32%	2.22% ± 0.43%
17	6.15% ± 1.34%	1.11% ± 0.32%	1.91% ± 0.43%
18	5.98% ± 1.33%	1.06% ± 0.29%	1.60% ± 0.42%
19	6.16% ± 1.38%	1.06% ± 0.29%	1.37% ± 0.37%
20	6.16% ± 1.34%	1.10% ± 0.31%	1.28% ± 0.32%

Table 5.7: Proportion of time scanning can be completed checking between 1 and 20 frequencies - random points, 95% confidence

Chapter 6

Using Time-of-Day and Location-based Mobility Profiles to Improve Scanning During Handovers in WiMAX

6.1 Introduction

When a WiMAX/802.16 subscriber mobile station (MS) wishes to join the network, it must follow the network entry procedure. This involves scanning for a frequency on a base station (BS). It is expected that the MS be required to perform repeated scanning to maintain connectivity to the network by moving from one BS to another while moving throughout the coverage area. This process of changing BSs is called a handover.

Mobile WiMAX networks are expected to provide support for all Internet applications. High bandwidth and realtime applications, in a network with high mobility are still important open areas of research. WiMAX is considered as a competitive technology to replace or enhance existing 3G cellular networks. One of the most important issues that needs improvement in order to maintain its competitiveness is handover performance when dealing with large volumes of real-time data in a highly mobile environment [24].

In our previous work [41], we introduced frequency scanning strategies used to reduce the time a MS spends finding an available channel of a WiMAX/802.16 BS. These basic strategies are now implemented within user mobility profiles.

In this Chapter, we introduce time-of-day and location-based mobility profiles to aid MSs during the scanning phase when performing a handover. If a mobile user often follows the same set of routes, the MS can exploit this to its advantage. We would like to capture the fact that a user travels to work in the morning, returns home in the evening or heads out to the gym on Wednesday nights as either time-of-day or location-based mobility profiles. Using these profiles, an MS reduces the number of

frequencies monitored in the scanning process during handovers.

6.1.1 Results of the Chapter

We propose two types of mobility profiles, *time-of-day* (TOD) and *location plus trajectory* (LPT), used by MSs in order to reduce the time spent searching for a frequency during handover scanning. We show a reduction in the time required for the scanning operation over our previous work [41]. Our strategies require no additional network support and only limited memory and computational resources of the MS.

We evaluated the use of TOD and LPT mobility patterns by modeling and simulating an area of WiMAX coverage with two MS user classes along with two different mobility models. In addition, we utilized real-world mobility trace data to evaluate the model.

The remainder of the chapter is organized as follows. In Section 6.2, we describe the WiMAX/802.16 network entry procedure and handovers. In Section 6.3, we briefly revisit our previous work in order to give a background to our new mobility profiles. In Section 6.4.1, we introduce the TOD user mobility profiles. In Section 6.4.2, we introduce LPT user mobility profiles. We briefly analyze the Manhattan mobility pattern in Section 6.5. We provide a description of the simulation environment along with the simulation results in Sections 6.6 and 6.8. Finally, we discuss ongoing work and conclude in Section 6.9.

6.2 Network Entry and Handovers in WiMAX

WiMAX/802.16 defines a network entry procedure for an MS wanting to establish a network connection via a BS. The MS must first scan to find a frequency in use by a BS. It does this by listening to each possible frequency until it hears the frame preamble. This takes a minimum of two frames [8, 104], at each frequency. After finding the frequency, it must synchronize with the BS by waiting for the Downlink Map (DL_MAP). The DL_MAP is a map of the timeslot locations in use for the frame. Once the MS has synchronized with the channel, it then must listen for the Downlink and Uplink Channel Descriptors (DCD and UCD) that are periodically sent by the BS. Then the MS must wait for a contention slot (determined from the UCD) in

order to perform Initial Ranging with the BS. Initial ranging is used to determine the transmit power level of the MS in order to reach the BS.

When performing a handover, the MS must essentially perform the steps of network entry with a neighbor of the currently serving BS. Since the network entry process has many steps, if we can improve upon the time it takes for one or more of these steps to complete we can provide an improved access time of service. In this chapter, we focus on handover *scanning*.

6.2.1 Scanning in WiMAX

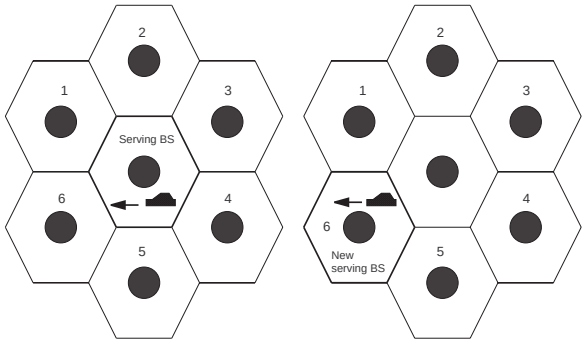
Scanning is an activity conducted by a MS. The goal of scanning is to acquire a downlink signal from a BS. Scanning is done by monitoring each possible frequency until a downlink signal is received. The exact number of available frequencies depends on the regulatory provisioned bandwidth (varies from one country to another), physical specification (several) and bandwidth per channel (several options available per physical specification). Scanning is performed during the initial network entry procedure and continues periodically to aid the MS in the selection of a suitable target BS for a handover to maintain network connectivity while in motion.

WiMAX/802.16 provides optional support for network assisted handovers where the BS currently serving the MS can obtain the information of neighboring BSs over the network. The serving BS periodically sends this information as a MOB NBR-ADV message to the MSs.

6.2.2 WiMAX Handovers

As an MS moves throughout the coverage area, maintaining connectivity is done via performing handovers between neighboring BSs. An example is shown in Figure 6.1, where an MS must choose one of six neighboring BSs, in this case neighbor six is chosen. Selection of the best handover target can be complex since the MS must scan for neighboring BSs to find a suitable target based on a number of criteria such as signal strength or error rates. Since a handover is an important function, an MS should perform the scanning and determine a target BS *before* beginning the handover. The IEEE 802.16 standard supports temporarily suspending the uplink and

downlink communication between the MS and BS in order to allow the MS to perform scanning for neighboring BSs. While communication is suspended, the data streams must be buffered on either side. Any improvement on the time it takes for the MS to complete its scanning operation improves the performance of the communications, reducing delays.



(a) MS before handover. (b) MS after handover.

Figure 6.1: Mobile station handover.

While WiMAX supports three kinds of handovers: hard handovers, Fast Base Station Switching (FBSS), and Macro Diversity Handover (MDHO), hard handovers is the only mandatory handover scheme [35]. Since a hard handover is a break then make mechanism, data being sent across a connection is more sensitive to delays in connecting to the new BS. Additionally, according to the WiMAX/802.16 specifications, handovers must be completed with a latency of less than 50ms [8].

6.3 Revisiting Scanning Strategies

In the following section, we briefly give a background on default scanning in WiMAX and our previous work.

6.3.1 Default WiMAX Strategy

The IEEE 802.16 standard [8] specifies that an MS must keep a nonvolatile storage where it stores the last set of operational parameters. When an MS wishes to acquire a downlink channel, it uses this stored information. Whenever an MS fails to obtain

the downlink channel, it then continuously scans the possible channels of the downlink frequencies until a time it finds a downlink signal.

6.3.2 Frequency Scanning Strategies

When an MS is turned on for the first time, all frequencies are equally likely since the MS has no history. That is to say if there are n frequencies, that the frequencies f_1, f_2, ... , f_n have an initial probability distribution $p(f_1), p(f_2), \ldots, p(f_n) = \frac{1}{n}$.

From the initial setup, since an MS has no previous history, it must simply start scanning frequencies in increasing order. As the MS performs a number of successful scanning operations, it can determine an *order of frequencies* from its observations. This history can be used to make scanning more efficient.

We previously proposed two different kinds of strategies to aid an MS in frequency selection for scanning. The first was a more general solution where an MS built a profile based on its history of successful scanning operations. This was based only upon the frequencies the MS had previously used. The second type of strategy leveraged the optional neighbor advertisements from a BS along with a profile built from the MSs history of handovers BS pairs.

Frequency Scanning Strategy

In our Most Frequently Used (MFU) strategy, a frequency of occurrence distribution over the radio frequencies is built using the history of the number of times each frequency is used. The frequency that is used the most has the highest probability. The MS keeps an ordered list starting with the most frequently used frequency. In future scanning operations, the MS scans the frequencies from the most frequently used to the least frequently used frequencies.

Previous Handover Strategy

In this strategy, we assume that the MS is provided with the list of all neighboring BSs and their frequencies in a MOB NBR-ADV message from its serving BS. Even though this information is provided to the MS, the MS must guess at which of the neighboring BSs it should attempt to perform a handover. The MS builds, for each

BS it visits, a most probable list of handover target neighbors based on its previous handover history. The MS uses this list, along with the MOB NBR-ADV message, to aid in selecting future handover target BSs.

6.4 Using Mobility Profile Strategies to Improve Scanning

In the following sections we describe our extended scanning and handover strategies with the introduction of two mobility profiles, *time-of-day* (TOD) and *location plus trajectory* (LPT).

6.4.1 Time-of-day Mobility Profile

We introduce our time-of-day (TOD) mobility profile. The MSs uses time-of-day information when building and maintaining their user mobility profile. Things that can be taking into consideration are:

1. Time of the day (morning, evening etc.).

2. Day of the week (weekday vs. weekends).

3. The current user, for multiple users.

For each time-of-day period, the MS keeps a list of *most probable* frequencies used during the period.

When an MS is about to perform a scanning operation, it looks up its time-of-day profile to determine the order of frequencies to check. Figure 6.2 shows a sample lookup table for two time-of-day profiles TOD_1 and TOD_2. For TOD_1, frequency f_2 is selected first, then f_1 while for TOD_2 the reverse is true. $TOD_{(x,y)}$ represents the y^{th} frequency of time-of-day period x.

We assume that the first time an MS is started the set of frequencies are of equal probability. Over time, the MS builds a set of mobility profiles, each corresponding to an orderings of the set of frequencies. The MS determines the number of mobility profiles to store based on its own time-of-day mobility patterns.

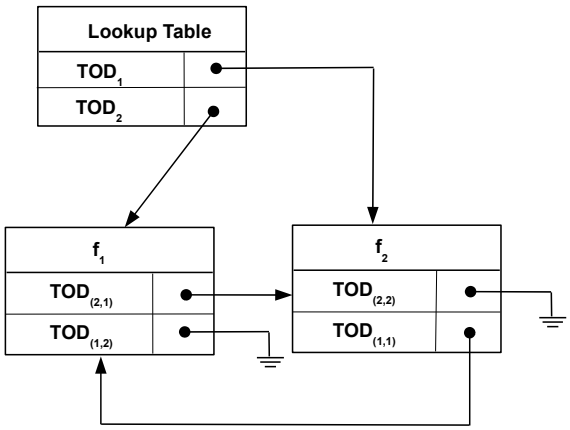


Figure 6.2: Mobile station lookup for time-of-day mobility profile frequency ordering.

The MS has a function that takes the time-of-day as a parameter and performs a table lookup for which time-of-day mobility profile to use. It then uses the *most probable* list of frequencies associated with this time-of-day profile.

6.4.2 Location Plus Trajectory Mobility Profile

We introduce our *location plus trajectory* (LPT) mobility profile. In the previous section, we presented a mobility profile that maintains lists of *most probable* frequencies depending upon the time of day or day of the week. This did not take into consideration information such as current location or direction of travel. In the LPT mobility profile, a MS equipped with a GPS or other means of localization uses this information when building its *most probable* frequency lists.

A MS uses its location as well as its trajectory to build a profile of several *most probable* frequency lists. The scene is depicted in Figure 6.3. We see MS A and MS B. Given an area of coverage, the region is divided into $n \times m$ zones. For each zone, the MS keeps a set of lists as a function of the trajectory of the MS within the zone.

As an example, Figure 6.4 shows how the MS determines which list from the current zone to use when performing scanning. In this example, the zone has four related lists, one for each quadrant Q1 through Q4. The MS determines which list to use by calculating which quadrant it is located in by using its current and previous

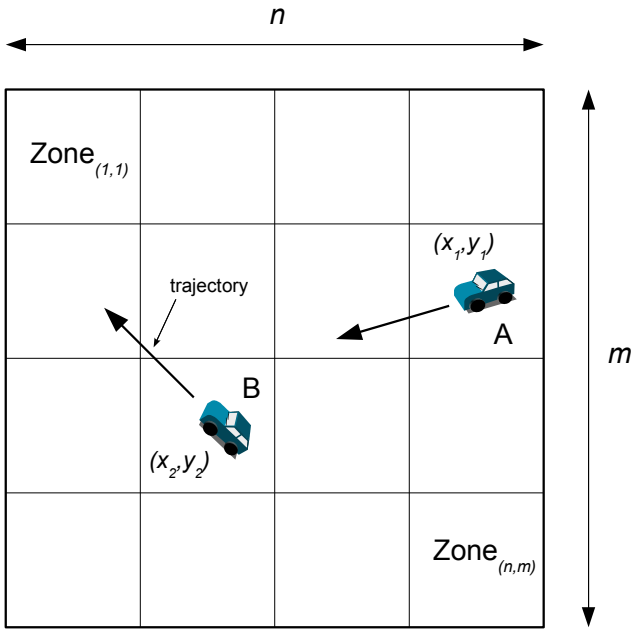


Figure 6.3: Overview of location plus trajectory.

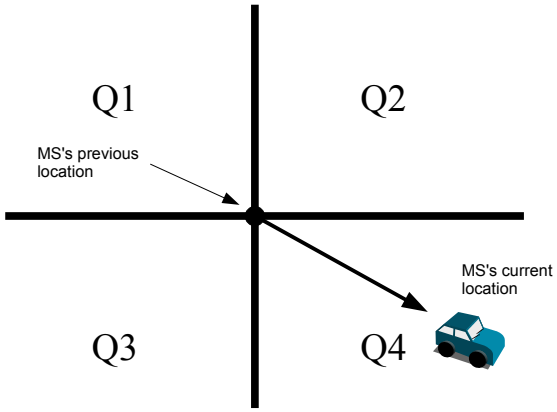


Figure 6.4: Determining a mobile station's trajectory.

locations. A lookup is performed in a similar fashion as the TOD mobility profiles shown in Figure 6.2, with the exception that each zone has multiple *most probable* lists as compared to the single TOD *most probable* list.

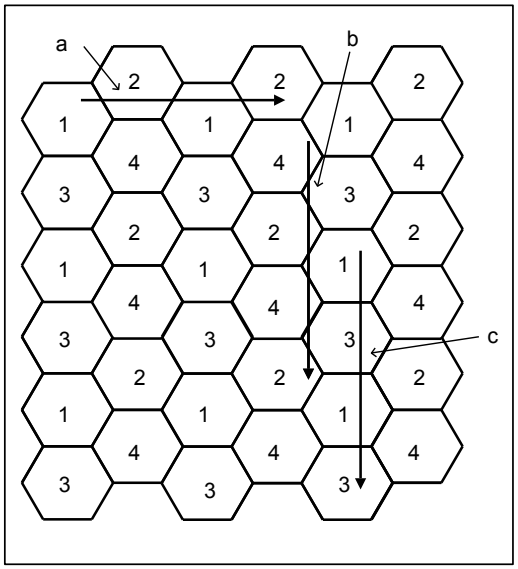


Figure 6.5: WiMAX coverage, frequency use, four cell clusters, Manhattan mobility model.

6.5 Analysis of Manhattan Mobility Model

To give some intuition on why maintaining a most probable list of frequencies to check during scanning as mobility profiles improves scanning time, we investigate the properties of a MSs path in the Manhattan model [115].

Refer to Figure 6.5 where we show a sample of three segments of possible MS paths and how the MS visits a set of BSs in the horizonal and vertical directions in the Manhattan model with a four cell clustering scheme. Cells labeled 1 through 4 are each assigned $\frac{1}{4}$ of the available frequencies. In the case of horizontal movement, the segment indicated by "a" in Figure 6.5, a MS visits a set of BSs that at most contain $\frac{1}{2}$ the total number of frequencies (BSs that fall in to cells numbered 1 and 2 of the four cell cluster).

In the vertical movement case, we see two cases of the set of BSs visited by a MS. The first is similar to the horizontal case and is depicted in Figure 6.5 the segment indicated by "c". The latter is shown in Figure 6.5, the segment indicated by "b",

Examine Figure 6.6 where we depict the area of a BS cell. We can calculate the area of the three sections and determine the probability of a user traveling, in the

vertical direction, through each part. Given a unit hexagon, the area of the triangle on the left hand side, with a base of $\sqrt{3}$ and a height of $\frac{1}{2}$ is $\frac{1}{4}\sqrt{3}\ units^2$. We multiply this by 2 in order to get the area of both triangles, giving us $\frac{1}{2}\sqrt{3}$. The area of the central part is $\sqrt{3}\ units^2$. The triangles represent $\frac{1}{3}$ of the area and the central part represents $\frac{2}{3}$ of the area of the hexagon.

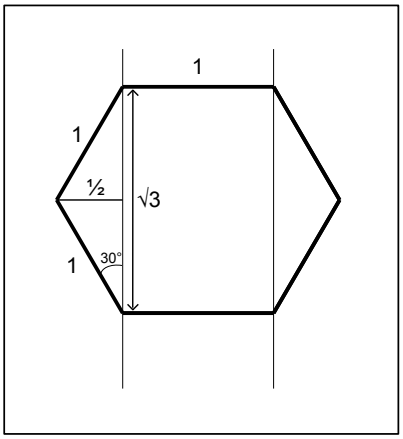


Figure 6.6: Calculating the area of a cell.

When moving in the vertical direction, the proportion of time a MS may use BSs that cover all of the available frequencies is only $\frac{1}{3}$ of the time (as with segment "b" in Figure 6.6) versus $\frac{2}{3}$ of the time when it only uses BSs that cover $\frac{1}{2}$ of the available frequencies (as with segment "c" in Figure 4).

In the Manhattan mobility model, this leads to the conclusion that with a higher probability, an MS repeatedly following similar paths is only exposed to $\frac{3}{4}$ of the frequencies in use in the area of coverage (assuming fixed frequency assignments). If an MS can check these frequencies first while scanning, it should gain an advantage.

6.6 Performance Evaluation

In order to evaluate the model, we first implemented a simulation to compare the WiMAX/802.16 default scanning strategy with the original *MFU* strategy and that of the TOD and LPT mobility profiles. Similarly, we implemented a second simulation

to compare the WiMAX default strategy, our original *previous handover strategy* and the TOD mobility profile for the previous handover strategy.

We first describe important parameters concerning the simulation: Network topology, metrics and measurements being used, the Mobile Station mobility patterns used, and time-of-day during the weekdays. Later we discuss the results of the simulations.

6.6.1 Network Topology

A ten kilometres by ten kilometres area is defined and covered by WiMAX/802.16 BSs with a one kilometre range. The base stations are positioned according to a cellular networking model [25] where each BS covers a hexagonal shaped region as shown in Figure 6.7. A total of 44 BS are used to cover the entire area.

Network Parameters	
Coverage Area	10 km x 10 km
No. Base Stations	44
Cell Radius	1 km
Cell Cluster Size	4
No. Frequencies	20
No. Channels	80
No. Channels per Cell	20
No. Frequencies per cell	5
TOD Periods	Weekday AM, Weekday PM
LPT Zones	16, 25
LPT Trajectory Lists	Quadrant based

Table 6.1: Simulation parameters.

The WiMAX/802.16 simulation parameters are chosen based on 100 MHz of bandwidth available for both uplink and downlink channels. Every channel is given 1.25 MHz of bandwidth for a total of 80 channels. A four cell reuse clustering scheme is implemented on the underlying 44 BSs. This setup has 20 channels available for each cell for the 80 channels available in total. We use 20 different frequencies, assigning five to each BS, each of which has four channels for MS connections. Two MS scanning times are implemented in the simulation scenario. Each simulation was run for a period of 100 weekday work days for 100 mobile stations.

In the model, the number of frequencies stored in memory is equal to the number of frequencies available in the coverage area. The simulation network parameters are summarized in Table 6.1.

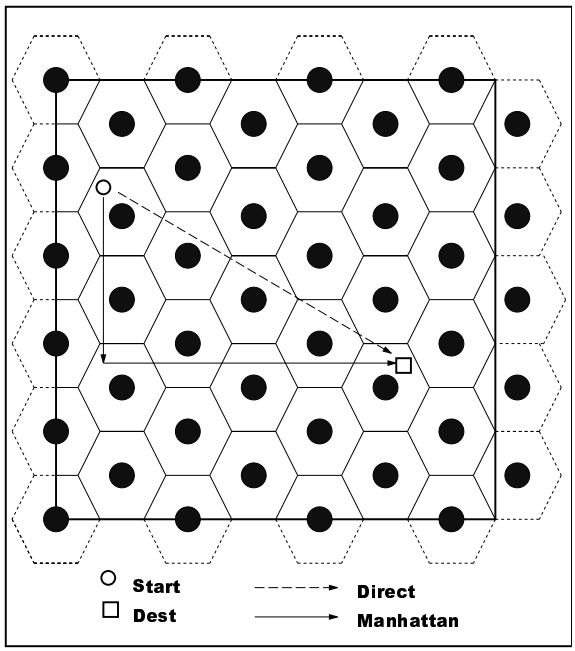


Figure 6.7: Simulation area with sample paths of a mobile station.

6.6.2 Metrics and Measurement

The purpose of this work is to improve upon (reduce) the number of frequencies an MS is required to check while scanning for a downlink connection to a BS. We measure the proportion of time that an MS can find the downlink with a certain number of frequencies checked. For the model and the frequency strategies, it is between one and 20 frequencies checked per scan. For the previous handover strategy, it is between one and six neighboring BSs. Ideally, we would like to have it so that 100% of the time an MS can successfully scan for a neighboring BS with only checking a single frequency, but the actual performance is dependent on the individual MSs mobility. From the observations of the simulations, we determine the expected number of frequencies required to be checked for a certain proportion of scanning operations. For example,

x% of the time, the MS can find a downlink channel with only scanning a single frequency. Results were obtained with a 95% level of confidence.

6.6.3 Mobile Station Mobility Models

For the simulations we used a variety of simulated and real-world mobility data. For the simulated mobility patterns we introduced two *classes* of user and two mobility models.

User Classes

1. Worker class - The worker mobility pattern models the scenario of a user that is commuting to and from work. The MS is given a set of random permanent *home* and *work* locations assigned within the area of WiMAX coverage. The worker travels from home to work and back once a day.

2. Wanderer class - The wanderer class is initially given a random start and destination location within the area of coverage. The wanderer moves towards its destination until it is reached. Then a new random destination is chosen and the wanderer continues moving towards this new destination. This repeats for the duration of the simulated day.

Mobility Models

The worker and wanderer classes are simulated using two different mobility models as shown in Figure 6.7. The first is the Manhattan model. The MS follows first the horizonal path, then the vertical path towards its destination. The second is the direct model, where the MS follows a direct, straight-line, course across the area of coverage to its destination. Since the wanderer class continuously selects a new destination when reaches the previous one, it is following the random waypoint model [81].

In order to study the impact of slight variances in the repeated mobility patterns, we also implemented a possibility of detour for the worker class. With some probability, the worker class user will detour from its path by up to 2 km. This was implemented for the Manhattan mobility model.

Real-World Mobility Traces

In addition to the simulated mobility patterns, we also used real-world mobility data from the Automatic Packet Reporting System (APRS) project [50]. Since the locations for each MS were only updated on the order of every few minutes, the intermediate locations between updates are interpolated along a direct path in order to simulate the mobility second by second.

6.6.4 Time-of-day Mobility Profiles

In order to reduce the complexity of the simulation, a naive set of time-of-day mobility profiles was implemented. The day is divided into two time periods, morning and evening, that have the ability to capture a worker class MSs commute to work in the morning in one time-of-day profile and its commute home in the second.

For the purpose of this work, we consider time-of-day parameters during the work week. The goal is to improve scanning times for users that typically repeat similar routes such as the daily commute to and from work. Therefore, we expect that the *worker class* will benefit the most from our strategies. We performed simulations for the WiMAX default, *MFU* and TOD frequency strategies in order to compare their results.

6.6.5 Location Plus Trajectory Mobility Profiles

In our evaluation of our LPT mobility profiles, we keep a total of four *most probable* lists for each of the $n \times m$ zones, as is depicted in Figure 6.4. We also let $n = m$ and tested two cases for the number of zones: 16 and 25. We would expect that as you increase the number of zones, there should be an increase in performance since the MS can build a more fine-grained mobility profile. We performed simulations for the WiMAX default, *MFU* and LPT frequency strategies in order to compare their results.

6.7 Simulation Environment

The simulations were written in C++. There are three main classes: *Simulation*, *MobileStation* and *BaseStation*. The *BaseStation* class maintains its location, the list of frequencies in use and the assignment of channels to MSs. The *MobileStation* class is responsible for its current location, home location, destination location, its velocity (home, destination and velocity are not used for APRS simulations), the locations of visited BSs and their neighbours, the *most probable* frequency list, the scanning mechanism and the metrics collected to evaluate the scanning performance.

The *Simulation* class is responsible for maintaining instances of *MobileStation* and *BaseStation* objects. The *Simulation* class initializes the BSs and MSs and then executes the main event loop that drives the simulation. The simplified main event loop is shown in Algorithm 1. For a specified number of days, for the number of time periods in each day, for each MSs, if the MS has not completed its movement for the day, a call to its *updateMobileStationLocation()* routine is made to advance the MS to it's next location.

Algorithm 1 Simulation Main Event Loop

```
1: while simDay ≤ SIM_DAYS do
2:   while simTime ≤ END_OF_DAY do
3:     for i = 1 to NUM_MS do
4:       if MS[i].getState() ≠ DONE then
5:         MS[i].updateMobileStationLocation()
6:       end if
7:     end for
8:     simTime++
9:   end while
10:  simDay++
11: end while
```

The MSs *updateMobileStationLocation()* routine for the TOD simulation is shown in Algorithm 2. Here the MS will either use real-work APRS data or simulated mobility data depending upon the value of $MOBILITY_PATTERN$. If using APRS trace data, the MS will retrieve the next APRS location. If the TOD is the morning period, *TOD1*, the MS is using one of $MANHATTAN$ or $DIRECT$ mobility models,

it will obtain its next location by calling either *moveTowardsWork()* or *moveDirectTowardsWork()* respectively. Similarily, if the TOD is the afternoon period, the MS will obtain its next location by calling one of the *moveTowardsHome()* or *moveDirectTowardsHome()* routines. If the return code, *rc*, was OK then the MS will make a call to its *updateScanningStrategies*() routine.

The *updateScanningStrategies*() routine is responsible for determining if there is about to be a handover, performing the scanning process and updating the metrics for the WiMAX default, MFU and TOD strategies.

Algorithm 2 Update Mobile Station Location (TOD)

Require: $MOBILITY_PATTERN(APRS, MANHATTAN, DIRECT)$

```
1: if MOBILITY_PATTERN == APRS then
2:    rc = getNextAPRSMobileStationLocation()
3: else if timeOfDayIs() == TOD1 then
4:    if MOBILITY_PATTERN == MANHATTAN then
5:       rc = moveTowardsWork()
6:    else
7:       rc = moveDirectTowardsWork()
8:    end if
9: else if timeOfDayIs() == TOD2 then
10:   if MOBILITY_PATTERN == MANHATTAN then
11:      rc = moveTowardsHome()
12:   else
13:      rc = moveDirectTowardsHome()
14:   end if
15: end if
16: if rc == OK then
17:   updateScanningStrategies()
18: end if
```

6.7.1 APRS Overview

Automatic Packet/Position Reporting System (APRS) is a real-time tactical digital communications protocol for exchanging information between a large number of stations covering a large (local) area [48, 50, 47]. APRS can be implemented using a simple hardware and software setup. As a MS moves about its environment, it repeatedly reports its location information via the AX.25 protocol. APRS differs from

traditional packet networks in the following ways:

- Integration of maps and other integrated displays for visualization;
- Uses a one-to-many protocol, updates in real-time;
- Uses digital repeaters so prior knowledge of network is not required; and
- Provides a worldwide real-time communication capability via the Internet.

The information reported by APRS includes the MS identifier, a timestamp, its current latitude and longitude, course, speed and altitude.

6.8 Simulation Results

The following sections show the simulation results of time-of-day mobility profiles with the frequency and previous handover strategies. We simulated the scanning operation when the MS is about to perform a handover. First we show the results for the simulated mobility patterns. Then we show the simulation results with the APRS mobility trace data. Finally, the complete set of frequency scanning tables, including 95% confidence intervals can be found at the end of the chapter.

6.8.1 Time-of-Day Frequency Strategies

In this section, we present the simulation results for scanning strategies using time-of-day mobility profiles. Figure 6.8 gives the results of the simulations for the worker class using both Manhattan and Direct mobility patterns. It shows the percentage of time that the MS succeeds in acquiring a channel with checking $1, 2, ..., 20$ frequencies. We see a significant improvement over our previous MFU strategy without the time-of-day mobility profiles.

The results are further summarized in Table 6.2. MSs that used the time-of-day mobility profiles acquired the downlink frequency 50.29% of the time with scanning only five frequencies. This is an increase over the strategy without time-of-day profiles which successfully acquired the frequency 43.73% of the time and the default WiMAX strategy which only succeeded 28.96% of the time with scanning only five frequencies.

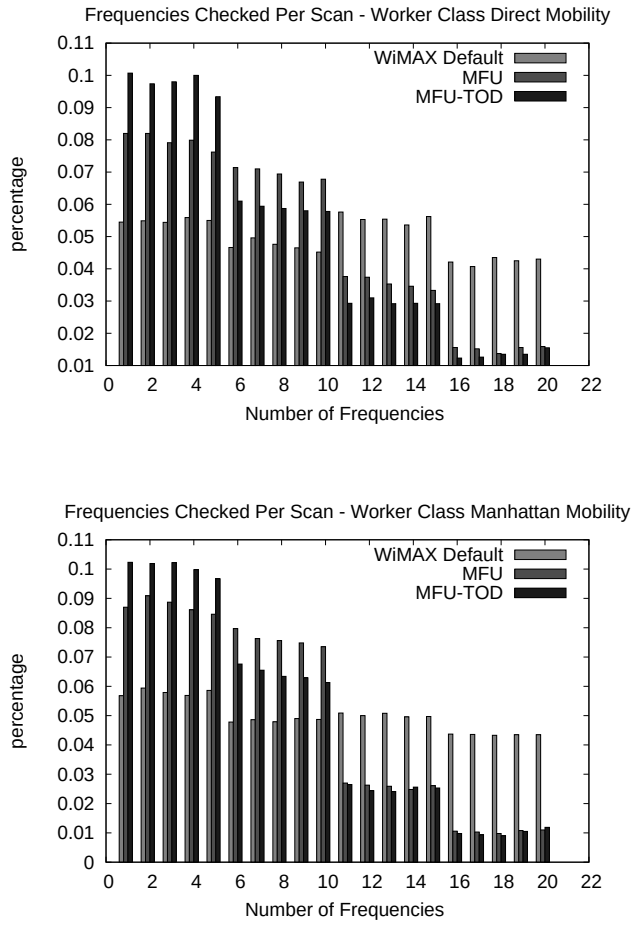


Figure 6.8: Frequencies checked per scan - Worker class.

Scanning Results		
	5 Freq	**10 Freq**
Default	28.96%	53.16%
MFU	43.73%	81.72%
MFU Time-of-Day	50.29%	82.36%

Table 6.2: Proportion of time scanning can be completed with checking 5, 10 frequencies - Worker class Manhattan mobility.

We then introduced the random detour to the worker class, Manhattan mobility, MSs daily commute. With a probability of 10%, a detour would be taken either on the way to work or on the way home from work with equal probability. The detour

could be up to 2km in distance from the MS's path. In Figure 6.9 we see that there is little impact on the performance of the strategy that uses time-of-day mobility profiles. The results are further summarized in Table 6.3.

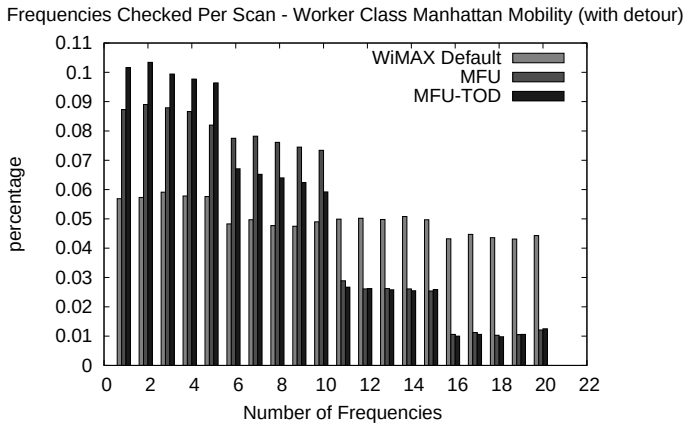


Figure 6.9: Frequencies checked per scan - effect of a random detour on Worker class with Manhattan mobility model.

Scanning Results		
	5 Freq	**10 Freq**
Default	28.87%	53.09%
MFU	43.28%	81.25%
MFU Time-of-Day	49.85%	81.64%

Table 6.3: Proportion of time scanning can be completed with checking 5, 10 frequencies - Worker class Manhattan mobility with detour.

Finally, since not all users may follow a predictable mobility pattern, we simulated the wanderer class and expected to see little advantage to building time-of-day user mobility profiles for this class of user. We also simulated a mixture of worker and wanderers. The topmost graph in Figure 6.10 shows the performance of our strategies with the wanderer class using the Manhattan mobility pattern. As expected, since the wanderer class has no predictable mobility pattern, there is no improvement in the number of frequencies checked during the scanning operation.

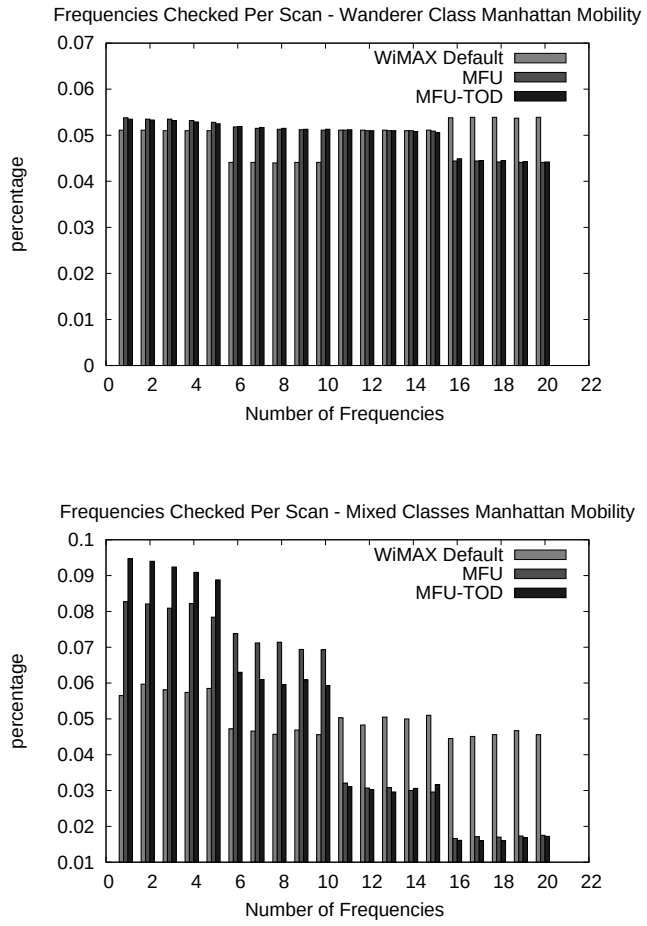


Figure 6.10: Frequencies checked per scan - Wanderer class and Mixed classes.

The bottommost graph of Figure 6.10 shows the results of the simulation when 80% of the MSs are workers and 20% are wanderers and they are further represented in Table 6.4.

Scanning Results		
	5 Freq	**10 Freq**
Default	29.02%	52.22%
MFU	40.63%	76.14%
MFU Time-of-Day	46.08%	76.46%

Table 6.4: Proportion of time scanning can be completed with checking 5, 10 frequencies - Mixed classes (80%/20%) Manhattan mobility.

6.8.2 Location Plus Trajectory Frequency Strategies

In this section, we present the simulation results for scanning strategies using time-of-day mobility profiles. Figure 6.11 gives the results of the simulations for the worker class using both Manhattan and Direct mobility patterns. It shows the percentage of time that the MS succeeds in acquiring a channel with checking $1, 2, ..., 20$ frequencies. We see a significant improvement over the previous MFU strategy and the TOD mobility profiles.

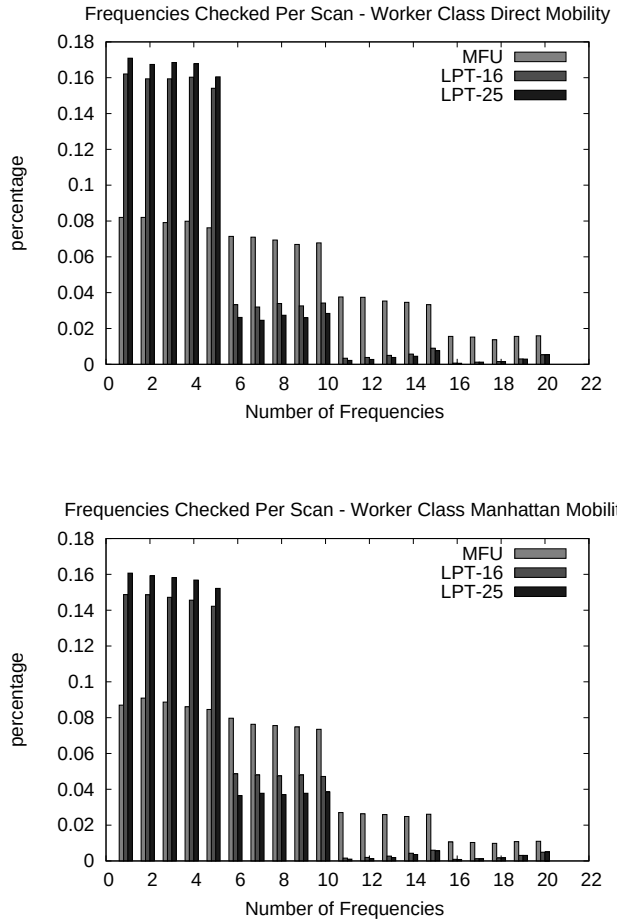


Figure 6.11: Frequencies checked per scan - Worker class.

The results for the worker class mobility are further summarized in Table 6.5. We observe that MSs using the LPT-16 and LPT-25 profiles acquired the downlink

frequency 73.23% and 78.71% of the time with scanning only five frequencies. This compares to 28.96% for WiMAX default, 43.73% for the original MFU strategy and 50.29% for the TOD mobility profile.

Scanning Results		
	5 Freq	**10 Freq**
Default	28.96%	53.16%
MFU	43.73%	81.72%
MFU Time-of-Day	50.29%	82.36%
LPT-16	73.23%	97.17%
LPT-25	78.71%	97.47%

Table 6.5: Proportion of time scanning can be completed with checking 5, 10 frequencies - worker class Manhattan mobility.

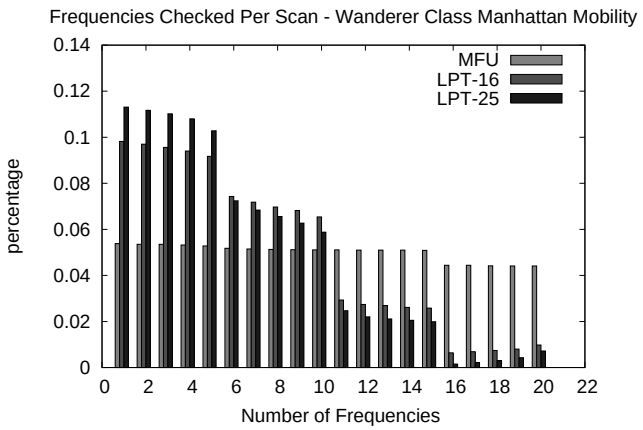


Figure 6.12: Frequencies checked per scan - Wanderer class - LPT mobility profiles.

Finally, as in the previous section, since not all users may follow a predictable mobility pattern, we simulated the wanderer class. Figure 6.12 shows the performance of our LPT strategies with the wanderer class using the Manhattan mobility pattern. In the case of the wanderer class, we do see improvement of using the LPT-16 and LPT-25 mobility profiles over the original MFU strategy. The LPT-16 and LPT-25 acquired the downlink frequency 47.65% and 54.58% of the time with scanning only five frequencies. This is an increase over the original MFU strategy which successfully

acquired the frequency 26.68% of the time within five frequencies. The results for the wanderer class mobility are summarized in Table 6.6.

It is worth noting that as expected, the LPT-25 mobility profile performed better when compared to the LPT-16 mobility profile. With an increased number of zones, each zone covered a smaller area thereby giving the MS a finer grained profile.

Scanning Results		
	5 Freq	**10 Freq**
MFU	26.68%	52.37%
LPT-16	47.65%	82.59%
LPT-25	54.58%	87.37%

Table 6.6: Proportion of time scanning can be completed with checking 5, 10 frequencies - Wanderer class Manhattan mobility.

6.8.3 Time-of-Day Previous Handover Strategy

In this section we present the simulation results for the previous handover strategy when using time-of-day mobility profiles. Recall that the previous handover strategy enables the MS using optional neighbor advertisement messages, sent by its currently serving BS, along with a mobility profile to make better choices during scanning. We have a simple set of time-of-day mobility profiles. We divide the day in to two time periods that have the ability to capture a worker class MSs commute to work in the morning in one time-of-day profile and its commute home in the second. In this simulation scenario, the WiMAX default was to simply scan the neighbors in the order given by the MOB_NBR-ADV messages.

Figure 6.13 shows the results of the simulations for the worker class using both Manhattan and direct mobility patterns. We see significant improvement over the previous strategy without the time-of-day mobility profiles.

The results are further shown in Table 6.9 where we see dramatic improvements in target BS selection when using time-of-day mobility profiles. This is intuitive since a worker class MS following a return path likely traverses the set of BSs in reverse order

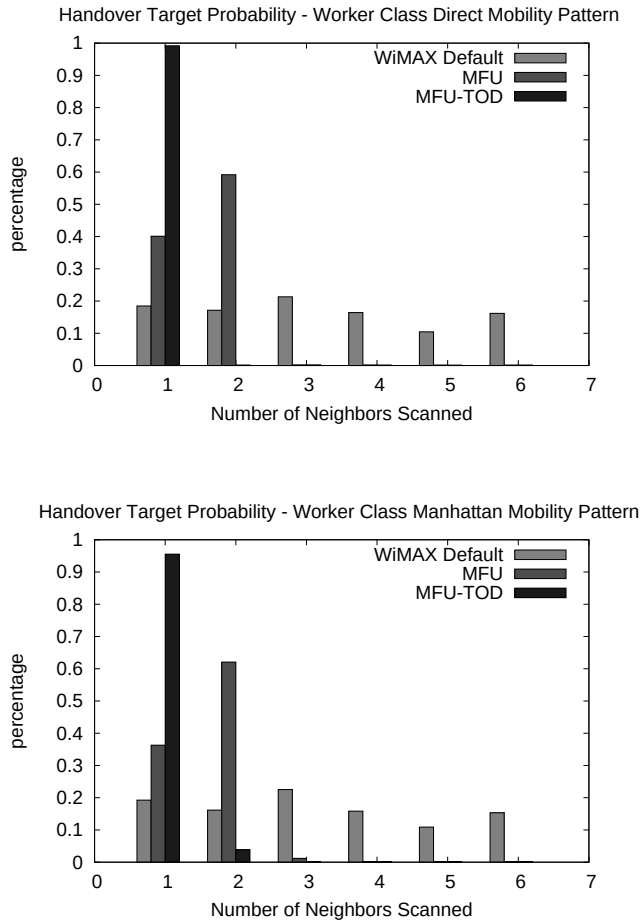


Figure 6.13: Base station neighbor scans per handover - Worker class.

Scanning Completed within Number of Neighbors			
No. Neighbors Scanned	WiMAX	MFU	TOD
1	19.26% ± 2.57%	36.29% ± 5.08%	95.54% ± 1.74%
2	16.15% ± 3.83%	62.07% ± 5.14%	3.82% ± 1.74%
3	22.54% ± 2.95%	1.19% ± 0.68%	0.2% ± 0.03%
4	15.83% ± 2.29%	0.16% ± 0.03%	0.17% ± 0.03%
5	10.9% ± 3.09%	0.16% ± 0.03%	0.13% ± 0.03%
6	15.31% ± 2.62%	0.14% ± 0.02%	0.15% ± 0.02%

Table 6.7: Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - Worker class Manhattan mobility, 95% confidence.

Scanning Completed within Number of Neighbors			
No. Neighbors Scanned	WiMAX	MFU	TOD
1	18.48% ± 3.21%	40.11% ± 5.38%	99.19% ± 0.03%
2	17.14% ± 3.7%	59.19% ± 5.37%	0.15% ± 0.03%
3	21.33% ± 3.42%	0.21% ± 0.03%	0.21% ± 0.03%
4	16.42% ± 2.87%	0.18% ± 0.03%	0.17% ± 0.03%
5	10.46% ± 2.71%	0.15% ± 0.03%	0.13% ± 0.03%
6	16.18% ± 3.33%	0.15% ± 0.03%	0.15% ± 0.03%

Table 6.8: Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - Worker class Direct mobility, 95% confidence.

Previous Handover Results		
	Manhattan	**Direct**
Default	19.26%	18.48%
MFU	36.29%	40.11%
MFU Time-of-Day	95.54%	99.19%

Table 6.9: Proportion of time handover target base station chosen on first attempt - Manhattan/Direct.

As in the previous section, we also studied the effect of a detour on the worker class MSs using the Manhattan mobility model. As before, with a 10% probability, the worker MS would make a detour of up to 2km away from its path during either its route to work, or its route home. The results are shown in Figure 6.14 and we see that there is minimal impact on the performance over the non-detour scenario.

Scanning Completed within Number of Neighbors			
No. Nbrs Scanned	WiMAX	MFU	TOD
1	19.30% ± 2.54%	50.06% ± 4.21%	95.35% ± 1.37%
2	16.09% ± 3.78%	47.54% ± 4.23%	3.39% ± 1.37%
3	22.52% ± 2.90%	1.52% ± 0.68%	0.42% ± 0.05%
4	15.93% ± 2.24%	0.33% ± 0.05%	0.31% ± 0.05%
5	10.78% ± 3.02%	0.27% ± 0.04%	0.23% ± 0.04%
6	15.37% ± 2.57%	0.28% ± 0.04%	0.31% ± 0.04%

Table 6.10: Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - Worker class Direct mobility with detour, 95% confidence.

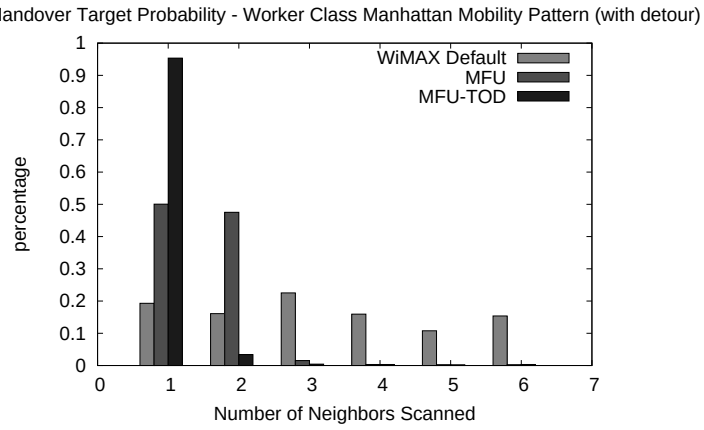


Figure 6.14: Base station neighbor scans per handover - effect of a random detour on Worker class with Manhattan mobility model.

Finally, as in the previous section, we simulated the wanderer class and expected to see little advantage to building user mobility profiles for this class of user. We also simulated a mixture of worker and wanderers. The topmost graph in Figure 6.15 shows the performance of the strategies with the wanderer class using the Manhattan mobility pattern. As expected, since the wanderer class has no predictable mobility pattern, there is no improvement in the scanning operation when selecting the target BS.

The bottommost graph of Figure 6.15 shows the results of the simulation when 80% of the MSs are workers and 20% are wanderers and they are further represented in Table 6.11.

Previous Handover Results	
	Manhattan
Default	18.34%
MFU	34.49%
MFU Time-of-Day	82.43%

Table 6.11: Proportion of time handover target base station chosen on first attempt - Mixed classes (80%/20%), Manhattan mobility.

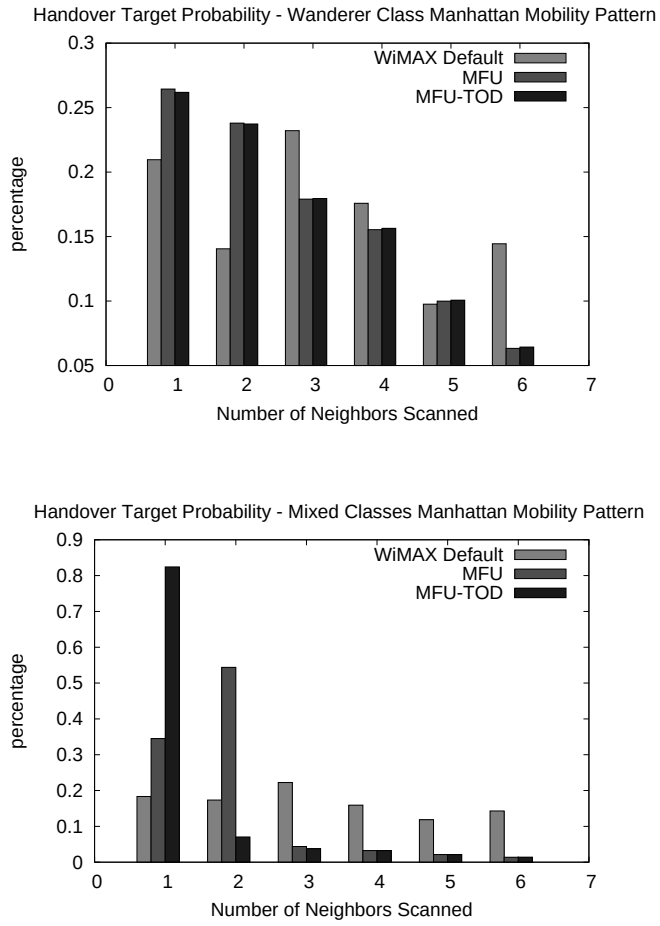


Figure 6.15: Base station neighbor scans per handover - Wanderer and Mixed classes.

6.8.4 Mobility Trace Data Simulation

In this section we present the results using mobility trace data from the APRS project. We performed a simulation with 35 MS mobility traces and our two *time-of-day* and *location plus trajectory* mobility profiles.

In Figure 6.16, we present the results for the *previous handover strategy* using the time-of-day mobility profiles. Here we do see an improvement of handover scanning frequency selection. When using time-of-day mobility profiles, the MS successfully acquires the correct BS target on the first attempt 63.25% of the time versus 54.19% for the previous mobility profile and only 17.64% for the WiMAX default strategy.

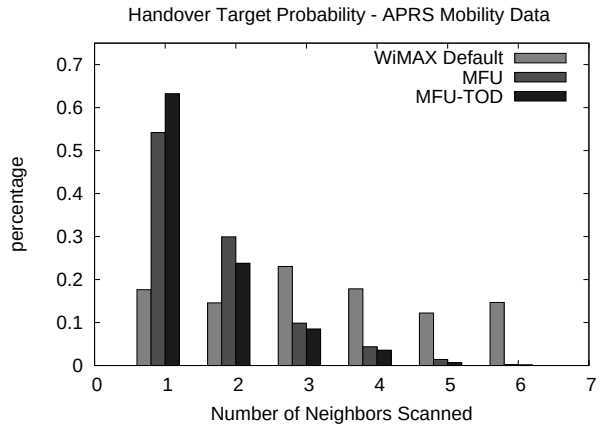


Figure 6.16: Base station neighbor scans per handover - APRS mobility data.

Scanning Completed within Number of Neighbors			
No. Nbrs Scanned	WiMAX	MFU	TOD
1	17.64% ± 3.07%	54.19% ± 2.63%	63.25% ± 3.77%
2	14.58% ± 1.97%	29.93% ± 1.73%	23.80% ± 2.32%
3	23.04% ± 2.16%	9.86% ± 1.46%	8.50% ± 1.48%
4	17.84% ± 2.69%	4.38% ± 0.94%	3.56% ± 0.88%
5	12.22% ± 1.78%	1.41% ± 0.55%	0.71% ± 0.35%
6	14.68% ± 2.34%	0.23% ± 0.17%	0.18% ± 0.16%

Table 6.12: Proportion of time scanning can be completed checking between 1 and 6 neighboring BSs - APRS mobility, 95% confidence.

Figure 6.17 shows the results of our frequency scanning strategies for LPT with 16 and 25 zones against our original MFU strategy using the APRS mobility trace data. We observe significant improvements through the use of the LPT mobility profiles.

The results are further summarized in Table 6.13. MSs that used the LPT-16 and LPT-25 mobility profiles acquired the downlink frequency 55.53% and 63.99% of the time with scanning only five frequencies. This is an increase over the MFU strategy which successfully acquired the frequency 32.43% of the time and the default WiMAX strategy which only succeeded 23.36% of the time with scanning only five frequencies.

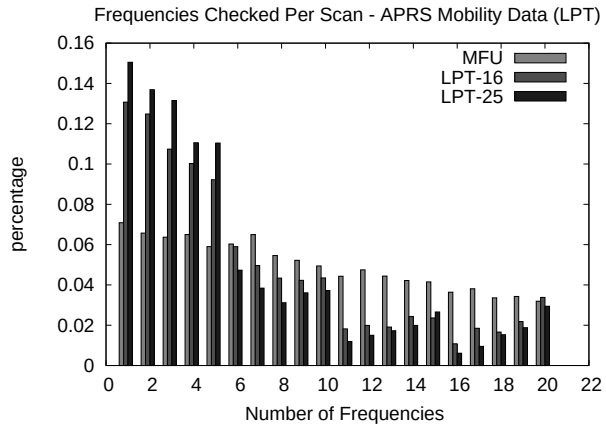


Figure 6.17: Frequencies checked per scan - APRS mobility data, LPT mobility profiles.

Scanning Results		
	5 Freq	**10 Freq**
Default	23.36%	46.53%
MFU	32.43%	60.58%
LPT 16 Zones	55.53%	79.31%
LPT 25 Zones	63.99%	83.01%

Table 6.13: Proportion of time handover scanning can be completed with checking 5, 10 frequencies - APRS mobility data.

6.9 Conclusion and Future Work

In our work, we have focused on improving the scanning times for WiMAX/802.16 MSs during handovers. We have introduced two mobility profiles, *time-of-day* and *location plus trajectory*, using our MFU scanning strategy, to aid in reducing the number of frequencies checked while scanning to find a downlink from a BS. We have provided some performance results from a set of experiments based on simulated and real-world mobility traces.

Through the simulated mobility experiments, we have shown that using either TOD or LPT mobility profiles improves handover scanning performance over that of the WiMAX/802.16 default and our previous work. For example, the worker class

using the Manhattan mobility model could complete scanning within five frequencies on average 50.29%, 73.23% and 78.71% of the time for the TOD, LPT-16 and LPT-25 mobility profiles respectively. This is an increase over the 43.73% seen from our original MFU and the 28.96% for the WiMAX default. Additionally, the LPT mobility profile provided significant improvement for the Wanderer mobility class where the TOD profile did not. We observed MSs completing scanning within five frequencies 47.65% and 54.58% of the time on average for LPT-16 and LPT-25 as compared to the 26.68% seen for our original MFU strategy. The Manhattan and Direct mobility patterns show two extremes of mobility. We would expect that in reality the mobility would be somewhere in between.

Using the real-world APRS mobility data, along with the LPT mobility profiles, the MS can complete scanning within five frequencies 55.53% and 63.99% of the time for LPT-16 and LPT-25 respectively. This is an improvement over the 32.43% for our original MFU strategy and the 23.36% for the WiMAX default strategy. We also show improvement in scanning performance when using the TOD mobility profile for the *previous handover* strategy when simulating with the APRS mobility trace data. However, a more sophisticated set of TOD mobility profiles should be implemented. In addition, other sets of mobility data should be examined, such as users of cellular telephones.

For the time-of-day mobility profiles, the MS is maintaining information on windows of times that it a) uses a certain frequency, and b) performs a handover from one BS to another. This window of time is not specific to any particular date, just a particular day or category of day of the week (i.e. Monday-Friday, Weekend). The profiles built with the initial frequency strategy will be of little other use since it only relates frequency use to a time period. With profiles for the handover strategies, we can extrapolate a general pattern of BSs traversed (i.e. user often travels through BS 1, BS 2, then BS 3 on Weekday mornings). This should not be a security or privacy issue since the information maintained in the profiles is rather coarse grained, is only maintained within the MS's internal memory and not shared with any other entity.

Future work includes providing a larger set of finer grained time-of-day mobility profiles. There is the additional problem of developing an MS algorithm for building

its optimal set of time-of-day profiles. The set of time periods that need to be profiled depends on the mobility patterns of individual MSs.

A further extension to the LPT model is to include building profiles for different routes an MS may take. The MS could then determine which route it is currently following and use this on which to base its scanning decisions. Additionally, the assignment of LPT zones to the coverage area could be done differently by examining the MSs routes and determining locations of roads etc.

Finally, an additional refinement to the model to investigate how different loads on the WiMAX/802.16 BSs, including multiple MSs attempting to perform ranging with the BS during the same ranging opportunities, may affect the scanning and ranging times. The work in this chapter is to be submitted for publication [42].

6.10 Overall Scanning Results

Scanning Results - Worker, Direct Mobility			
No. Freq Checked	WiMAX	MFU	TOD
1	5.45% ± 0.68%	8.20% ± 0.45%	10.07% ± 1.01%
2	5.49% ± 0.66%	8.20% ± 0.45%	9.74% ± 0.87%
3	5.44% ± 0.67%	7.91% ± 0.41%	9.80% ± 0.86%
4	5.59% ± 0.69%	7.99% ± 0.42%	10.00% ± 0.97%
5	5.50% ± 0.68%	7.62% ± 0.42%	9.34% ± 0.84%
6	4.66% ± 0.70%	7.14% ± 0.46%	6.10% ± 0.65%
7	4.96% ± 0.77%	7.10% ± 0.46%	5.94% ± 0.61%
8	4.76% ± 0.73%	6.94% ± 0.43%	5.87% ± 0.61%
9	4.65% ± 0.71%	6.69% ± 0.40%	5.80% ± 0.59%
10	4.52% ± 0.67%	6.78% ± 0.49%	5.78% ± 0.59%
11	5.76% ± 0.75%	3.76% ± 0.51%	2.93% ± 0.50%
12	5.53% ± 0.66%	3.74% ± 0.49%	3.10% ± 0.53%
13	5.54% ± 0.67%	3.53% ± 0.46%	2.92% ± 0.49%
14	5.36% ± 0.67%	3.46% ± 0.44%	2.93% ± 0.47%
15	5.62% ± 0.68%	3.33% ± 0.41%	2.92% ± 0.42%
16	4.21% ± 0.68%	1.56% ± 0.38%	1.23% ± 0.34%
17	4.07% ± 0.65%	1.52% ± 0.37%	1.26% ± 0.34%
18	4.35% ± 0.71%	1.37% ± 0.33%	1.35% ± 0.35%
19	4.25% ± 0.68%	1.56% ± 0.35%	1.35% ± 0.31%
20	4.30% ± 0.68%	1.59% ± 0.34%	1.55% ± 0.32%

Table 6.14: Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Direct mobility, 95% confidence.

Scanning Results - Worker, Manhattan Mobility			
No. Freq Checked	**WiMAX**	**MFU**	**TOD**
1	5.68% ± 0.74%	8.70% ± 0.36%	10.23% ± 0.76%
2	5.94% ± 0.77%	9.09% ± 0.38%	10.19% ± 0.76%
3	5.79% ± 0.73%	8.87% ± 0.38%	10.22% ± 0.75%
4	5.69% ± 0.73%	8.61% ± 0.41%	9.98% ± 0.75%
5	5.86% ± 0.75%	8.46% ± 0.37%	9.67% ± 0.78%
6	4.78% ± 0.77%	7.97% ± 0.38%	6.76% ± 0.56%
7	4.86% ± 0.80%	7.63% ± 0.41%	6.55% ± 0.54%
8	4.79% ± 0.77%	7.56% ± 0.43%	6.34% ± 0.56%
9	4.90% ± 0.80%	7.48% ± 0.38%	6.29% ± 0.53%
10	4.87% ± 0.78%	7.35% ± 0.47%	6.13% ± 0.49%
11	5.09% ± 0.81%	2.70% ± 0.39%	2.65% ± 0.41%
12	5.00% ± 0.79%	2.63% ± 0.39%	2.44% ± 0.36%
13	5.08% ± 0.81%	2.59% ± 0.37%	2.41% ± 0.36%
14	4.96% ± 0.79%	2.48% ± 0.35%	2.56% ± 0.37%
15	4.97% ± 0.79%	2.61% ± 0.35%	2.53% ± 0.34%
16	4.37% ± 0.70%	1.06% ± 0.34%	0.98% ± 0.32%
17	4.36% ± 0.72%	1.03% ± 0.31%	0.93% ± 0.30%
18	4.33% ± 0.71%	0.98% ± 0.31%	0.90% ± 0.28%
19	4.35% ± 0.70%	1.08% ± 0.32%	1.05% ± 0.29%
20	4.35% ± 0.71%	1.10% ± 0.30%	1.19% ± 0.32%

Table 6.15: Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Manhattan mobility, 95% confidence.

Scanning Results - Worker, Direct, Detour			
No. Freq Checked	WiMAX	MFU	TOD
1	5.69% ± 0.72%	8.73% ± 0.35%	10.16% ± 0.73%
2	5.73% ± 0.72%	8.9% ± 0.36%	10.34% ± 0.83%
3	5.91% ± 0.77%	8.79% ± 0.39%	9.94% ± 0.72%
4	5.78% ± 0.74%	8.66% ± 0.39%	9.77% ± 0.71%
5	5.76% ± 0.73%	8.2% ± 0.36%	9.64% ± 0.73%
6	4.83% ± 0.77%	7.75% ± 0.37%	6.71% ± 0.58%
7	4.97% ± 0.78%	7.82% ± 0.39%	6.52% ± 0.56%
8	4.77% ± 0.75%	7.61% ± 0.41%	6.4% ± 0.54%
9	4.75% ± 0.75%	7.45% ± 0.43%	6.24% ± 0.52%
10	4.9% ± 0.79%	7.34% ± 0.42%	5.92% ± 0.51%
11	4.99% ± 0.79%	2.89% ± 0.39%	2.67% ± 0.39%
12	5.02% ± 0.79%	2.61% ± 0.37%	2.62% ± 0.38%
13	4.98% ± 0.77%	2.62% ± 0.36%	2.58% ± 0.37%
14	5.08% ± 0.8%	2.61% ± 0.34%	2.55% ± 0.36%
15	4.97% ± 0.79%	2.54% ± 0.33%	2.59% ± 0.35%
16	4.32% ± 0.69%	1.06% ± 0.31%	1% ± 0.31%
17	4.47% ± 0.72%	1.12% ± 0.32%	1.06% ± 0.31%
18	4.36% ± 0.69%	1.03% ± 0.3%	0.98% ± 0.28%
19	4.31% ± 0.7%	1.05% ± 0.29%	1.06% ± 0.29%
20	4.43% ± 0.7%	1.21% ± 0.3%	1.25% ± 0.29%

Table 6.16: Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Direct mobility with Detour, 95% confidence.

Scanning Results - Worker, Manhattan Mobility				
No. Freq	WiMAX	MFU	LPT-16	LPT-25
1	5.68% ± 0.74%	8.70% ± 0.36%	14.87% ± 0.71%	16.07% ± 0.70%
2	5.94% ± 0.77%	9.09% ± 0.38%	14.86% ± 0.70%	15.92% ± 0.66%
3	5.79% ± 0.73%	8.87% ± 0.38%	14.72% ± 0.73%	15.82% ± 0.69%
4	5.69% ± 0.73%	8.61% ± 0.41%	14.56% ± 0.72%	15.68% ± 0.70%
5	5.86% ± 0.75%	8.46% ± 0.37%	14.22% ± 0.68%	15.22% ± 0.65%
6	4.78% ± 0.77%	7.97% ± 0.38%	4.87% ± 0.66%	3.65% ± 0.64%
7	4.86% ± 0.80%	7.63% ± 0.41%	4.80% ± 0.66%	3.78% ± 0.64%
8	4.79% ± 0.77%	7.56% ± 0.43%	4.75% ± 0.63%	3.70% ± 0.61%
9	4.90% ± 0.80%	7.48% ± 0.38%	4.80% ± 0.64%	3.78% ± 0.62%
10	4.87% ± 0.78%	7.35% ± 0.47%	4.72% ± 0.61%	3.85% ± 0.59%
11	5.09% ± 0.81%	2.70% ± 0.39%	0.16% ± 0.08%	0.10% ± 0.05%
12	5.00% ± 0.79%	2.63% ± 0.39%	0.20% ± 0.09%	0.13% ± 0.04%
13	5.08% ± 0.81%	2.59% ± 0.37%	0.27% ± 0.09%	0.19% ± 0.05%
14	4.96% ± 0.79%	2.48% ± 0.35%	0.43% ± 0.11%	0.36% ± 0.09%
15	4.97% ± 0.79%	2.61% ± 0.35%	0.60% ± 0.13%	0.56% ± 0.11%
16	4.37% ± 0.70%	1.06% ± 0.34%	0.09% ± 0.02%	0.08% ± 0.02%
17	4.36% ± 0.72%	1.03% ± 0.31%	0.12% ± 0.03%	0.12% ± 0.03%
18	4.33% ± 0.71%	0.98% ± 0.31%	0.17% ± 0.04%	0.17% ± 0.04%
19	4.35% ± 0.70%	1.08% ± 0.32%	0.31% ± 0.06%	0.31% ± 0.06%
20	4.35% ± 0.71%	1.10% ± 0.30%	0.49% ± 0.09%	0.50% ± 0.09%

Table 6.17: Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Manhattan mobility, 95% confidence.

Scanning Results - Worker class, Direct Mobility				
No. Freq	WiMAX	MFU	LPT-16	LPT-25
1	5.45% ± 0.68%	8.20% ± 0.45%	16.21% ± 0.86%	17.09% ± 0.74%
2	5.49% ± 0.66%	8.20% ± 0.45%	15.94% ± 0.72%	16.74% ± 0.64%
3	5.44% ± 0.67%	7.91% ± 0.41%	15.94% ± 0.74%	16.85% ± 0.67%
4	5.59% ± 0.69%	7.99% ± 0.42%	16.03% ± 0.75%	16.78% ± 0.66%
5	5.50% ± 0.68%	7.62% ± 0.42%	15.41% ± 0.71%	16.05% ± 0.66%
6	4.66% ± 0.70%	7.14% ± 0.46%	3.33% ± 0.66%	2.62% ± 0.58%
7	4.96% ± 0.77%	7.10% ± 0.46%	3.20% ± 0.62%	2.46% ± 0.53%
8	4.76% ± 0.73%	6.94% ± 0.43%	3.39% ± 0.64%	2.74% ± 0.59%
9	4.65% ± 0.71%	6.69% ± 0.40%	3.26% ± 0.59%	2.60% ± 0.54%
10	4.52% ± 0.67%	6.78% ± 0.49%	3.42% ± 0.58%	2.84% ± 0.53%
11	5.76% ± 0.75%	3.76% ± 0.51%	0.34% ± 0.17%	0.22% ± 0.12%
12	5.53% ± 0.66%	3.74% ± 0.49%	0.38% ± 0.17%	0.27% ± 0.12%
13	5.54% ± 0.67%	3.53% ± 0.46%	0.50% ± 0.18%	0.38% ± 0.15%
14	5.36% ± 0.67%	3.46% ± 0.44%	0.57% ± 0.19%	0.45% ± 0.13%
15	5.62% ± 0.68%	3.33% ± 0.41%	0.90% ± 0.21%	0.76% ± 0.14%
16	4.21% ± 0.68%	1.56% ± 0.38%	0.07% ± 0.02%	0.06% ± 0.02%
17	4.07% ± 0.65%	1.52% ± 0.37%	0.12% ± 0.03%	0.12% ± 0.03%
18	4.35% ± 0.71%	1.37% ± 0.33%	0.15% ± 0.04%	0.15% ± 0.04%
19	4.25% ± 0.68%	1.56% ± 0.35%	0.30% ± 0.06%	0.29% ± 0.06%
20	4.30% ± 0.68%	1.59% ± 0.34%	0.53% ± 0.10%	0.53% ± 0.09%

Table 6.18: Proportion of time scanning can be completed checking between 1 and 20 frequencies - Worker class, Direct mobility, 95% confidence.

Scanning Results - APRS Mobility Data				
No. Freq	WiMAX	MFU	LPT-16	LPT-25
1	4.35% ± 0.81%	7.09% ± 0.73%	13.07% ± 1.20%	15.05% ± 1.03%
2	4.59% ± 0.65%	6.57% ± 0.54%	12.48% ± 0.91%	13.69% ± 0.94%
3	4.70% ± 0.75%	6.37% ± 0.72%	10.74% ± 0.95%	13.15% ± 0.98%
4	5.51% ± 0.74%	6.50% ± 0.78%	10.02% ± 1.01%	11.06% ± 0.97%
5	4.21% ± 0.61%	5.90% ± 0.63%	9.22% ± 1.07%	11.04% ± 1.23%
6	5.09% ± 0.63%	6.03% ± 0.60%	5.90% ± 0.57%	4.73% ± 0.60%
7	4.29% ± 0.69%	6.50% ± 0.83%	4.96% ± 0.64%	3.84% ± 0.68%
8	4.53% ± 0.72%	5.46% ± 0.59%	4.34% ± 0.59%	3.12% ± 0.55%
9	4.47% ± 0.58%	5.22% ± 0.72%	4.23% ± 0.60%	3.61% ± 0.63%
10	4.79% ± 0.73%	4.94% ± 0.69%	4.35% ± 0.59%	3.72% ± 0.67%
11	5.05% ± 0.75%	4.43% ± 0.54%	1.82% ± 0.49%	1.19% ± 0.36%
12	5.13% ± 0.62%	4.75% ± 0.49%	1.99% ± 0.40%	1.50% ± 0.49%
13	4.86% ± 0.74%	4.44% ± 0.67%	1.91% ± 0.38%	1.73% ± 0.48%
14	5.03% ± 0.72%	4.22% ± 0.57%	2.43% ± 0.61%	1.99% ± 0.46%
15	5.32% ± 0.77%	4.15% ± 0.65%	2.36% ± 0.52%	2.66% ± 0.77%
16	5.14% ± 0.70%	3.64% ± 0.65%	1.08% ± 0.25%	0.61% ± 0.33%
17	6.09% ± 0.78%	3.81% ± 0.58%	1.85% ± 0.63%	0.95% ± 0.37%
18	5.75% ± 0.84%	3.36% ± 0.50%	1.66% ± 0.40%	1.53% ± 0.47%
19	4.98% ± 0.73%	3.43% ± 0.57%	2.18% ± 0.38%	1.88% ± 0.35%
20	6.13% ± 1.18%	3.19% ± 0.50%	3.38% ± 0.96%	2.94% ± 0.85%

Table 6.19: Proportion of time scanning can be completed checking between 1 and 20 frequencies - APRS mobility data, 95% confidence.

Chapter 7

Prioritized Access for Emergency Stations in Next Generation Broadband Wireless Networks

7.1 Introduction

Historically, wireless communications for emergency services predate the modern cellular network by almost fifty years. The first one-way police radios were introduced by the Detroit Police Department in 1928. The first two-way systems were introduced in New Jersey in 1933 [69]. Since this time, dedicated emergency communication systems have been developed separate from the cellular networks. With the recent introduction of packet-based broadband wireless networks supporting a wide variety of applications, such as WiMAX/802.16, it is pertinent to investigate the integration of emergency communication systems with public wireless networks.

Various standardization bodies [6, 15] define the four following kinds of emergency communications:

1. Citizen to Authority - Citizen's communications with the authorities such as placing a 911 call.
2. Authority to Citizen - Authority's communications with citizens, such as an early warning system.
3. Authority to Authority - Authorities' ability to communicate amongst themselves, including between different agencies.
4. Citizen to Citizen - Citizens' ability to communicate with family and friends in time of crisis.

When emergency situations arise, wireless cellular networks can be overloaded with huge increases in the number of users attempting to gain access. In [19], the authors indicate that during an emergency situation the network can experience as

much as a tenfold increase over normal network demand. This can lead to life-threatening conditions if authorized emergency users cannot maintain service with the network.

In Orthogonal Frequency Division Multiple Access (OFDMA) based WiMAX networks, the network entry process begins with the MSs sending a random Code Division Multiple Access (CDMA) ranging code during a contention period after synchronizing with a BS. This process is anonymous, meaning a BS must process all ranging requests in order to determine if an MS can be granted entry.

The motivation of this work is twofold. First, we attempt to reduce or eliminate the interference between emergency and non-emergency MSs during the CDMA initial ranging contention period. Second, we attempt to have a BS process ranging requests from emergency MSs ahead of non-emergency MSs wherever possible. This requires the ability to distinguish between emergency and non-emergency MSs. We study a single uplink channel in use by a WiMAX/802.16 BS.

While the focus of our work here is on the initial ranging operation, the same principles can be applied to the periodic ranging, handover ranging and bandwidth request operations.

7.1.1 Results of the Chapter

In this chapter, we propose strategies that can enable BSs in a WiMAX/802.16 network to provide prioritized network entry access to an emergency class of MS. Our proposed strategies are designed to reduce or eliminate the interference between emergency and non-emergency MSs during the contention-based CDMA initial ranging process. They can also let a BS determine the type of MS attempting initial ranging and gives a BS flexibility on controlling which MSs are permitted to continue the network entry process even before the actual ID of the MSs are known. We evaluate the performance of our strategies through a series of simulations of the CDMA ranging operation when a BS is hit by a sudden burst of MS ranging messages.

The remainder of the chapter is organized as follows. In Section 7.2, we give a background to the problem and some previous work. In Section 7.3 we introduce our proposed strategies for prioritizing access to emergency MSs. In Section 7.4

we analyze the potential collisions between MSs during the CDMA initial ranging process. We provide a description of the simulation environment along with the simulation results in Section 7.5. Finally, we discuss ongoing work and conclude in Section 7.6.

7.2 Background

Initial ranging is an important step in the WiMAX/802.16 network entry process of a MS. It is used by a MS to determine the transmit power and timing offsets, in order to synchronize transmissions, to gain access to the network via a BS.

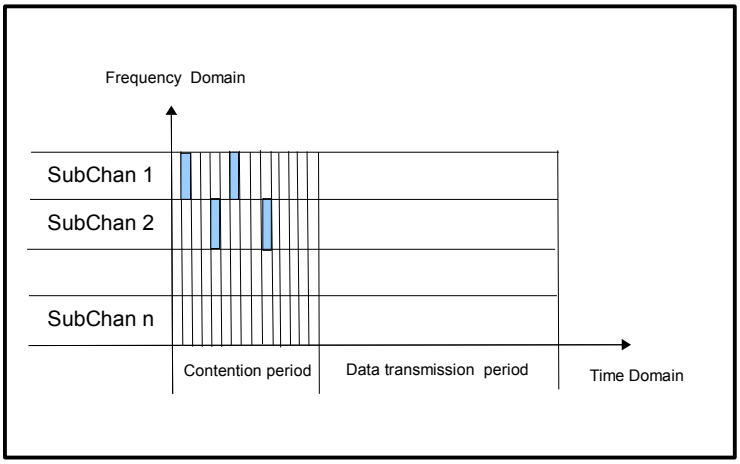


Figure 7.1: Contention and data period of OFDMA frame.

In OFDMA based WiMAX/802.16, a MS must first synchronize with the downlink (DL) and uplink (UL) via the DL-MAP and UL-MAP messages. Figure 7.1 shows the contention and data transmission periods of an OFDMA frame. The contention periods are structured into contention slots. From the UL-MAP, an MS obtains the ranging opportunities in terms of the number of contention slots in the next frame as well as the CDMA codes to be used. An example of the ranging opportunities for a frame is shown in Figure 7.2. One or more groups of six subchannels are allocated for the ranging process. The CDMA codes are a set of special pseudonoise (PN) 144-bit numbers. There are 256 codes available divided into the following types, initial ranging, periodic ranging, handover ranging and bandwidth requests.

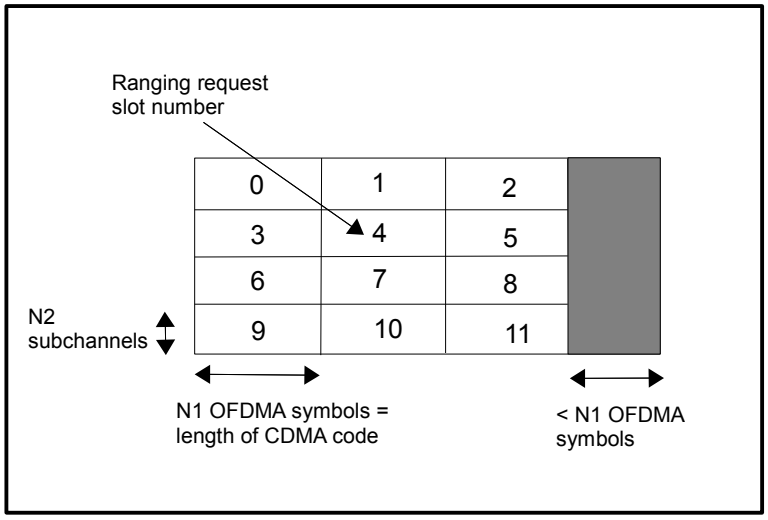


Figure 7.2: Ranging opportunities in contention period.

The MS randomly chooses a ranging slot within its initial contention window and a PN CDMA ranging code that it transmits twice in two consecutive contention slots as shown in Figure 7.3. The MS watches for the number of available contention slots advertised in each frame and then transmits its chosen CDMA ranging code during an advertised contention period.

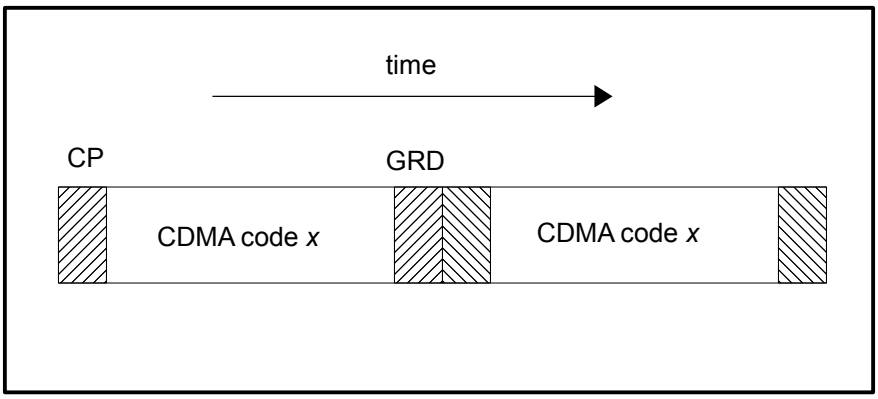


Figure 7.3: CDMA code sent over two consecutive contention slots.

When the MS sends its ranging request, it sets a timer (T3). When T3 expires and the MS has not heard a response from a BS it assumes its request was lost or not heard. The MS then enters the backoff phase, doubling its initial contention window

size and selecting a random opportunity within its new contention window size. This process repeats until the MS hears a response from a BS or a maximum number of retries has been reached.

The BS responds to the MS in a future downlink frame by using the CDMA code sent by the MS indicating on which contention slot the message was received to identify the MS. Since the MS only sends a random CDMA code, a BS has no knowledge of whether the MS is an emergency or non-emergency station. This causes problems, since there is no way to prioritize access to stations until further in the network entry process of establishing a connection with a BS. This leads to emergency and non-emergency stations interfering with and competing on the same contention slot/code pairs.

7.2.1 Previous Work

Several works focus on the ability of citizens to make emergency calls to authorities via 911 has been undertaken [7, 14, 2]. While this is important, our focus is on providing support for *authority to authority* communications.

The authors in [95] discuss the current state of the standardization activities for emergency communications as communications are starting to migrate to the Internet. They emphasize that there are a large number of organizations involved and it is difficult to clarify and coordinate all the different requirements.

For the remainder of this chapter, we will first discuss wireless communication systems designed for use in emergency scenarios. They can be grouped into two main categories, ad hoc based and infrastructure based. We then discuss related work of interest to emergency communication systems.

Ad hoc Network Based Emergency Architecture

Ad hoc networks are networks where all "nodes" are both senders of data and forwarders of data. The nodes collaborate to form a network and data is forwarded through the network in a multi-hop fashion. Due to their lack of infrastructure, ease of configuration and distributed routing protocols, they are suited well for emergency communications where there is no infrastructure or the infrastructure has been

destroyed. The following are a few specific ad hoc networks for use in emergency situations.

The Enhanced Communications Scheme Combining Centralized and Ad hoc Networks (ECCA) [70] is a hybrid architecture designed to collect damage assessment information during times of disaster by maintaining connections between BSs and ad hoc nodes via paths created via multiple hops. In the case where nodes lose connection to a BS, they communicate with neighboring BSs via other nearby nodes. Routes are discovered and maintained by eavesdropping on neighbor node broadcasts.

The Integrated Cellular and Ad hoc Relaying System (iCAR) [110] describes a new system that integrates cellular and ad hoc networks. iCar can reduce congestion by the introduction of Ad hoc Relaying Stations (ARS) used to balance traffic and relay messages from on BS to another in a dynamic fashion.

The MOBIDIS project [17] provided the design and implementation of a middleware platform in order to support distributed systems in MANETs. The goal of the project is to provide a multi-layer management system for MANETs. It requires that nodes have not only a wireless interface, but also special hardware in order to determine distances from its neighbors. A Predictive Layer uses a probabilistic algorithm to determine future connections of nodes.

The WIreless DEployable Network System (WIDENS) [21] is an ad hoc network that is optimized for public safety applications. The focus of WIDENS is on the integration of different technologies under a common platform. It allows user communications and sensor data collection without the use of fixed infrastructure.

FleetNet [84] designs a ad hoc network for vehicle-to-vehicle communications. Fleetnet forms scatternets interconnecting isolated pockets of an ad hoc network through the use of roadside wireless Internet gateways. These Internet gateways fill in the gaps of the network, acting like clusterheads, where vehicles in one pocket can send messages to vehicles in another via the gateways.

Finally, the National Institute of Standards and Technology (NIST) has developed a distributed testbed for first responders (DTFR) [18]. NIST works with various agencies in order to identify first responder needs in order to develop standards

and technologies for interoperable wireless communications. The DTFR was developed after the events of September 11, 2001 where different agencies were unable to communicate with each other, greatly hindering the coordinated emergency response. The testbed provided a mechanism for first responders to exchange information using a self-organizing MANET consisting of PDAs using WiFi links. The network can use any PDA in range to aid in relaying messages.

Infrastructure Based Emergency Architecture

The integration of IEEE 802.11 and 3G networks [51] attempts to give the best of both worlds. While 3G networks provide low speed connections they are available in most locations. The ability to seamlessly hop between the cellular and much faster, shorter range IEEE 802.11 hotspots gives the system more flexibility.

The MIRAI system [79] is designed to enable the seamless integration of multiple wireless technologies. The system provides functionality for network selection, paging and handovers between different networks.

TErrestrial Trunked RAdio (TETRA) [20] is a specialized radio communications system that is used by emergency services. TETRA is an ETSI standard. It uses lower frequencies than cellular networks, giving greater coverage areas. TETRA BSs have the ability to work alone providing service to local MSs. MSs that are within range of a BS can also relay messages for neighbor MSs not in range of the BS.

AVAYA [4] is a mobile communications unit that designed to be deployed in order to provide emergency communications. It interoperates with the public telephone network, cellular, wireless and satellite networks. It uses the advantages of available infrastructure, only reverting to satellite communications when all else fails.

An architecture of geographically distributed ticket servers [33] has been proposed. It can issue tickets to traffic flows identifying the priority that should be given based on the context of the current network conditions. This way, users of higher importance, such as emergency personnel, would be given a higher priority.

Other Related Work

The 3rd Generation Partnership Project is active in defining telecommunication standards. In their work item, [2], they discuss requirements for emergency calls to be routed to an emergency centre. Their work in [1] states that while investigation is underway to provide standardized Priority Service for emergency responders in circuit switched speech communications, there is a need to undertake this effort for packet-based (e.g. IP) networks. The IEEE 802 group has also recently released a *Call for Interest* [93] on creating a study group for emergency service provisioning.

DSRC/WAVE, which stands for Dedicated Short Range Communication/Wireless Access in Vehicular Environments, is currently under development as part of the IEEE 802.11p standard [12]. DSRC/WAVE describes an ad hoc system for intervehicular and vehicle to road side communications, typically in the range of a few hundred metres.

Bandwidth allocation and admission control are two important aspects of a wireless network that require more attention [98]. The radio resources are limited and must be allocated in an efficient manner to ensure QoS requirements are met. Admission control must be applied in order to limit the number of devices connected to the network in order to avoid overwhelming the system. The authors in [98] provide some queueing analysis and a non-cooperative game theory approach to the problem of bandwidth allocation and admission control.

The authors in [89] describe an enhancement for medium access and radio resource management for ad hoc networks based on UTRA time division duplexing (TDD). It is based on the Fleetnet system designed for inter-vehicular communications. They reserve a flexible number of communication slots for *high priority* communications, but state that they expect that these high priority messages will be rare in practice.

The authors in [114] propose an adaptive call admission control protocol for wireless networks where they dynamically increase and decrease the number of reserved *guard channels* in order to provide increased handover performance. The authors in [88] propose another adaptive call admission control algorithm for QoS/revenue optimization for CDMA cellular systems. Their admission control is based on bandwidth degradation. They computing a bandwidth degradation threshold by which

to admit a new call or handover. Their work is focused on "make-before-break" or "soft" handovers where the handover is intra-system and intra-frequency.

Another recent work is an enhancement to the cellular network giving priority to emergency calls over non-emergency calls [116]. Their focus was to reduce the impact of preemption on non-emergency calls. In their work, emergency calls studied were *citizen to authority* communications and they did not discuss emergency responders. They also assume that emergency calls will not be made while the user is in motion.

The authors in [101] provide a scheme to ensure priority medium access for emergency messages in an ad hoc network. This is done by sending out of band pulses on the control channel while sending emergency data. All nodes are required to monitor the control channel at all times and can determine when neighbors are sending emergency messages.

There have been a number of works on priority service for 3G cellular networks including priority access schemes in [59], a common packet channel access scheme [82] and a priority stack random access scheme for Wideband Code Division Multiple Access (W-CDMA) [23].

The authors in [112] discuss some of the requirements on IP networks for emergency services. They focus primarily on the communication of *citizen to authority* aspect of emergency services, more specifically on providing location information of a 911 VoIP caller to the authorities. The call may be in a fixed LAN or attached to a Wireless-LAN (WLAN). XML is used for encoding location information.

The optimal placement of a single WiMAX/802.16j relay stations has been studied in [87]. The area of coverage was a grid with an arbitrary coordinate system. Also, the SSs are assumed to be fixed with stable links to the RS, their long term resource requirements are known in advance and can be used in the RS placement process. This kind of scenario would be good for general network planning, but would not be applicable for emergency services. The authors in [107] propose a scheme to optimize the efficiency of WiMAX/802.16 RSs. They propose new protocol mechanisms based on MAC PDU concatenation and MAC SDU aggregation to reduce the wasted resources when the data to be transmitted does not use the entire DL or UL burst. The authors in [106] propose a path selection algorithm for MSs in WiMAX/802.16j

networks. Path information is stored at the MR-BS and the path selection algorithm is based upon the expected link throughput computed by each RS of the link between itself and its downstream node.

Most of the previous work on OFDMA WiMAX/802.16 initial ranging focused on either determining the optimal number of contention slots [76, 118] or on maximizing the signal detection at the BS [90, 113]. None is related to emergency services.

The authors in [94] propose a solution for providing two levels of quality of service, high and low priority, in WiFi/IEEE 802.11 networks by dividing the contention window in half. They had fixed designated high and low priority zones, but determined an overlap function to allow low priority traffic to contend on a portion of the high priority contention zone when traffic load is light.

We describe some previous work focusing on prioritizing emergency voice calls in cellular networks. Public Use Reservation with Queuing All Calls (PURQ-AC) [19] has been proposed. In PURQ-AC, separate buffers are maintained for emergency and public incoming calls. As resources are made available, the buffers are served in a round-robin fashion with the emergency buffer given one fourth of the allocation. The authors in [34, 118, 116] look at prioritizing emergency calls through preemption or delay of public calls. Their focus is on analyzing the load on the system while prioritizing emergency traffic and minimizing disruption to the non-emergency traffic.

Summary

There has been little analysis on an integrated, dual use (emergency and public) broadband wireless communication system supporting multimedia applications. One thing that is missing is a detailed study on the performance of a single system incorporating both public and emergency communications.

It was seen in the aftermath of the attacks of September 11, 2001, that the cellular networks were overwhelmed by too many users trying to access the network simultaneously. In the context of emergency communications, there is a need to find a way to avoid the interference from too many users. We must limit entry to the network to authorized emergency users. New adaptive access control schemes are required.

Our work focuses on WiMAX/802.16 data network access. We focus on the earliest step of the MS network entry process, initial ranging. We are aiming to prioritize emergency access by reducing the interference between emergency and non-emergency MSs during the CDMA initial ranging process with a BS as well as attempting to process emergency ranging requests ahead of non-emergency requests whenever possible. Even if there is no collision between emergency and non-emergency MSs, a BS must still process all ranging requests since CDMA initial ranging is an anonymous process.

7.3 Proposed Strategies for Prioritized Emergency Access

We propose a set of strategies for prioritizing emergency access to an OFDMA based mobile WiMAX/802.16 network. We envision a collaborative system where neighboring BSs coordinate on-the-fly to best serve the set of MSs attempting to gain access to the network. The BSs can adjust system parameters to meet real-time network conditions. BSs implement strategies such as adjusting the contention window for emergency MSs, advertise special emergency CDMA ranging codes, or delay sending range response messages to non-emergency MSs.

Initially, a BS advertises a single set of contention opportunities for initial ranging. BSs will treat all MSs as equals with no special prioritization. As the BSs detect events occurring in various locations in the network, the strategies are applied locally or globally and can be adjusted to best handle the changing situation on the ground. This facilitates providing a priority to emergency MSs network access. The detection of events impacting the network are beyond the scope of this initial work.

In the following sections, we introduce our strategies used by the BSs in order to facilitate the emergency prioritized access including the use of a sliding emergency contention window, a new class of CDMA ranging codes and allowing BSs to delay responses to non-emergency MSs for a portion of their ranging timeout.

7.3.1 Emergency Contention Zone

Our first strategy is to provide a dedicated portion of the ranging contention period to emergency MSs. This can be done in various ways. The first is to divide the

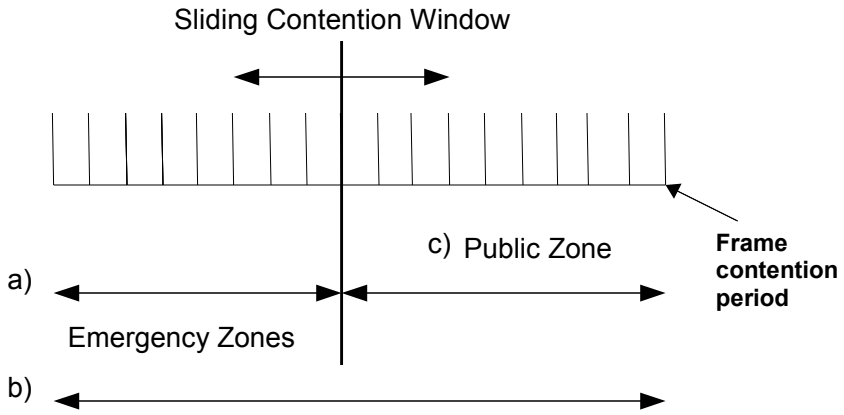


Figure 7.4: Split and overlapping contention windows for emergency and public MSs.

ranging contention period into separate windows, one for emergency MSs and one for non-emergency MSs. The second way is to extend the first idea by splitting the contention period into separate windows, but allowing the emergency MSs to utilize all slots within the contention period while limiting the non-emergency MSs to their subset of slots. Figure 7.4 depicts the contention period available and the sliding contention window to break up the contention period into emergency and public zones. Alternately, whole ranging contention periods for one frame or multiple sequential frames can be allocated to emergency MSs, thus delaying the contention of non-emergency MSs.

In order for MSs to determine when they can perform ranging, WiMAX/ 802.16 BSs advertise the initial ranging contention period with the CDMA Initial Ranging Information Element (IE) of the UL-MAP. To support the split contention window, we introduce a new CDMA Emergency Ranging IE message. This Emergency_Ranging_IE message informs emergency MSs of the dedicated emergency ranging slots allocated by a BS during the next frame.

With the first and third strategies, a BS can know the number of emergency MSs attempting to enter the network by monitoring the emergency ranging contention slots. With the second strategy, the total number of emergency MSs cannot be known since MSs pick a random CDMA code from the pool of available codes. Only

the emergency MSs that transmit their CDMA code during the emergency contention window are known since the process is anonymous. We address the anonymous CDMA codes in the following section.

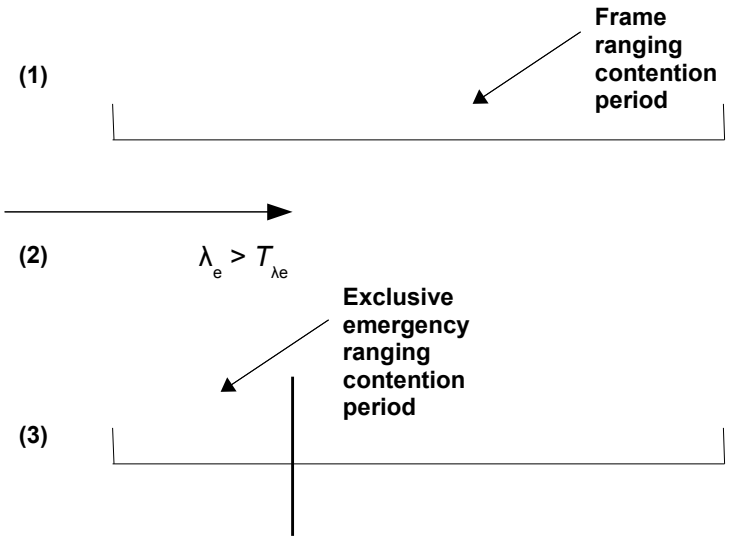


Figure 7.5: Threshold of arrivals of emergency MS ranging attempts and emergency contention zone.

The size of the emergency contention zone is determined by some threshold of emergency events such as the number of emergency MSs currently connected to the BS or the arrival rate of initial ranging messages from emergency MSs. Figure 7.5 shows the process of determining the size of the emergency contention zone. Initially, a BS advertises a single contention period available for ranging. This is depicted in Figure 7.5 (1). As the arrival rate of emergency range requests, λ_e crosses a threshold T_{λ_e}, as shown in Figure 7.5 (2), it triggers the BS to determine an exclusive contention period for emergency ranging in future frames as shown in Figure 7.5 (3). As an alternative, the BS can use a threshold of the number of emergency MSs currently connected to make this decision.

By setting the emergency zone to 100% of the contention region and not advertising public initial ranging, each BS can revert to an emergency only mode and know

that all range requests received are from emergency MSs.

7.3.2 Emergency CDMA Ranging Codes

WiMAX/802.16 defines a series of CDMA codes to be used for ranging. The MS randomly chooses one of the codes and sends it during a random ranging contention slot. If a BS hears the code, it sends a reply to the MS with the code and slot used. This is an anonymous process since the code used is not related to the ID of any particular MS.

In order to break this anonymity at the BS, we introduce a new emergency category of CDMA ranging code. By introducing a new type of ranging code, a BS can determine which MSs are emergency stations and should be given a higher priority. A portion of the total number of ranging codes are designated for emergency MSs only. Now, a BS can determine exactly which MSs are emergency devices, although not their IDs, and make decisions accordingly. The BS determines the breakdown on the number of codes to assign to emergency and non-emergency MSs based on the changing conditions of the network in a similar way as for determining the size of the emergency contention zones.

7.3.3 Base Station Delayed Response

Once a BS has the ability to distinguish between non-emergency and emergency MSs, either through non-overlapping contention zones or the use of special CDMA codes, it can then use this knowledge to make further decisions regarding MSs that are performing initial ranging.

When an MS sends its CDMA ranging code it sets a timer, T3, which is 60ms by default. We propose that a BS, under conditions of stress due to the presence of emergency MSs, delays responses to any non-emergency MS for a portion of the T3 timer value while waiting to see if any new emergency MSs initial ranging requests arrive. This delay can be determined in a similar manner as for setting the size of the emergency contention zone and the number emergency CDMA codes.

7.4 Analysis of Mobile Station Collisions

We present our analysis of the collision model to show the expected number of collisions between MSs in the overlapping contention zones as described in Section 7.3.1. We have the following parameters:

- the number of contention slots, s;
- the number of ranging codes, c;
- the number of slot/code pairs, S;
- the number of MSs, n;
- the number of MSs that select a given slot/code pair, r

Given that we have s contention slots and c CDMA codes, the probability that a given MS selects a given slot/code pair is denoted by

$$\frac{1}{(s \times c)} = \frac{1}{S}$$

Since we are looking to examine the number of collisions this is further expanded to determine the probability that a given set of r out of n MSs select the same slot/code pair as

$$\left(\frac{1}{S}\right)^r \left(1 - \frac{1}{S}\right)^{n-r}$$

From this, we can calculate the probability that *any* set of r MSs select the same slot/code pair as

$$\binom{n}{r} \left(\frac{1}{S}\right)^r \left(1 - \frac{1}{S}\right)^{n-r}$$

Continuing, calculate the expected number of slot/code pairs having a collision of r users as

$$S\binom{n}{r}\left(\frac{1}{S}\right)^{r}\left(1-\frac{1}{S}\right)^{n-r}$$

Then, we calculate the expected number of collided users across all slot/code pairs for a given value of r as

$$rS\binom{n}{r}\left(\frac{1}{S}\right)^{r}\left(1-\frac{1}{S}\right)^{n-r}$$

And finally, we obtain the expected number of collided users across all slot/code pairs when $r > 1$ (there is a collision when more than one MS select a given slot/code pair) as

$$\sum_{r=2}^{n} rS\binom{n}{r}\left(\frac{1}{S}\right)^{r}\left(1-\frac{1}{S}\right)^{n-r} \tag{7.1}$$

Figure 7.6 shows the expected number of collided users in the overlapping region of the contention period for the WiMAX default operation as well as with a dedicated emergency contention zone covering 25%, 50% and 75% of the contention period. The calculations are made using Equation 7.1 for 10/25, 25/50, 50/75 and 50/100 (emergency/non-emergency) MSs attempting ranging on a single frame with 16 contention periods and 128 CDMA ranging codes.

With these proposed strategies, it is possible for the authorities to switch the system into a completely emergency mode that eliminates all attempts at ranging by non-emergency MSs.

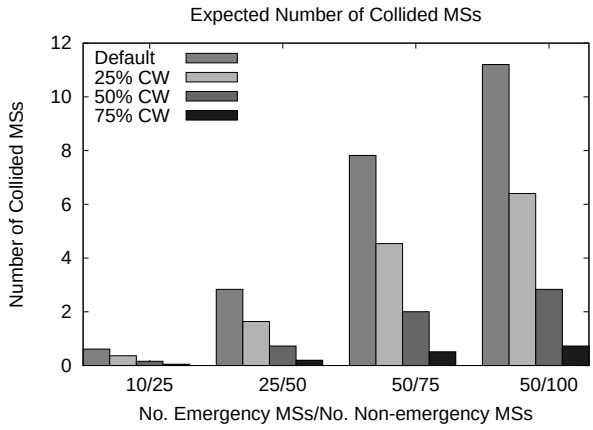


Figure 7.6: Expected number of MS involved in collisions, overlapping contention window.

7.5 Performance Evaluation

We developed a series of simulations to examine the performance of the CDMA ranging operation when a BS is hit by a sudden burst of MS ranging arrivals. We model the contention ranging process on a single frequency in use at a BS. An MS sends its randomly chosen CDMA ranging code over two consecutive ranging opportunities. If a BS hears an MSs request, this does not determine that it will be granted access. If the number of MSs attempting to gain access to the network resources is greater than the number that can be accommodated, then some MSs will not be granted access. The goal of the system is to provide for prioritized emergency access. We want to reduce, or eliminate interference between emergency and non-emergency MSs and to enable the BSs to process requests from emergency MSs ahead of those from non-emergency MSs.

Table 7.1 shows the simulation parameters tested. We conducted a series of simulations with a combination of emergency and non-emergency MSs competing for network access during the ranging contention period. In each simulation, the number of MSs arriving during each frame for five consecutive frames were varied from 10-50 and 25-100 for emergency and non-emergency MSs respectively. The initial contention window backoff was set to 16 slots for all MSs and there were 16 ranging opportunities

in each frame. The size of the emergency contention zone was varied between 0% and 75% of the total number of opportunities. A total of 128 CDMA codes were assigned for initial ranging. For each set of fixed parameters, the simulation was run for a series of 1000 trials and results taken. We measured two important metrics: (1) the probability of collisions between emergency and non-emergency (where possible) and (2) order of processing of MSs by the BS. All results were calculated with a 95% level of confidence.

Simulation Parameters	
Number of Emergency MSs per frame	10-50
Number of Non-emergency MSs per frame	25-100
Initial Contention Window Size	16
Ranging Opportunities per Frame	16
Emergency Zone Size (%)	0%-75%
Number of CDMA Ranging Codes	128
Emergency Codes (%)	0%-75%
Frame Length	5ms
Default T3 Timer	60ms
BS Delayed Response	10-30ms

Table 7.1: Simulation parameters.

7.5.1 Emergency Contention Zone

In the first scenario, we tested the emergency contention zone as described in Figure 7.4 a) where emergency MSs are given a portion of the contention ranging opportunities for their exclusive use.

This has the advantage that there is no chance of a collision between emergency and non-emergency MSs. However, the potential drawback is that if the emergency contention zone is too small in comparison to the initial contention window of the MS, the MS may be forced to wait a number of frames before sending it's CDMA code.

In order to evaluate our strategies we observe the percentage of emergency MSs that are processed by the BS within the first half of all MSs attempting ranging. The results are shown in Figure 7.7. The figure shows the results for the WiMAX default, 25%, 50% and 75% dedicated emergency contention zone scenarios. Here we

see that a separate contention window performs ahead of WiMAX default for and emergency contention zone of 50% and 75%, but not as well for the case of 25%. For the WiMAX default case, we see results of 49.4% to 49.6% of emergency MSs processed within the first 50% of all MSs. Similarly, we see between 40.05% to 41.98%, 57.60% to 62.65% and 71.24% to 75.84% of emergency MSs processed for a separate emergency contention window of 25%, 50% and 75% respectively. The results are summarized in Table 7.2

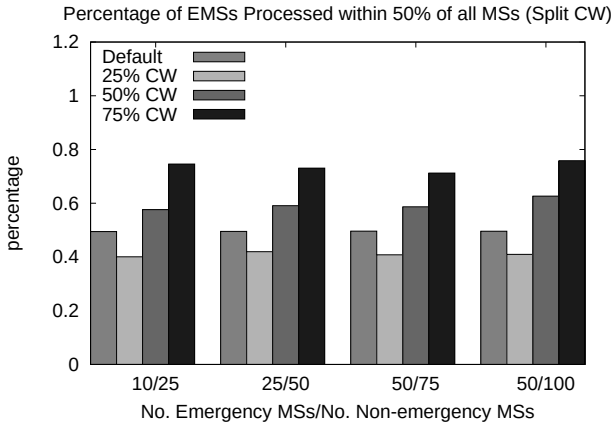


Figure 7.7: Percentage of emergency MSs processed in the first 50% of all MSs, split contention window.

Percentage of Emergency MSs Processed	
WiMAX Default	49.44% ±0.18% to 49.61% ± 0.08%
25% Emergency CW	40.05% ±0.27% to 41.98% ± 0.18%
50% Emergency CW	57.60% ±0.34% to 62.65% ± 0.13%
75% Emergency CW	71.24% ±0.12% to 75.84% ± 0.20%

Table 7.2: Percentage of emergency MSs processed within 50% of all MSs - split contention window.

For another measure, we also set a fixed value for the number of MSs that a BS allow entry on the current channel to 50 MSs. We then compare the number of emergency MSs granted entry as a percentage of the maximum possible. The results are shown in Figure 7.8. The WiMAX default values range from 28.14% to 38.24% of the maximum possible. With 50% and 75% emergency contention zone sizes, we see

improvements to 33.26% to 56.67% and 53.51% to 79.41% of the maximum possible respectively.

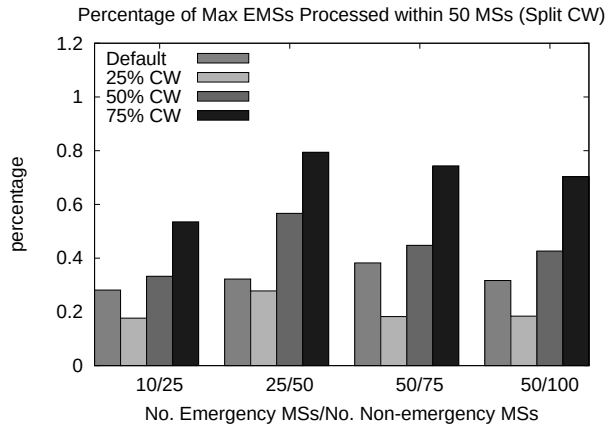


Figure 7.8: Percentage of maximum possible emergency MSs processed in the first 50 MSs, split contention window.

The next scenario tested is an extension of the exclusive emergency contention zone. As shown in Figure 7.4 b), the BS assigns a portion of the contention slots to the exclusive use of emergency MSs, but emergency MSs can contend on any of the contention slots advertised in the frame. This gives a portion of non-interfering emergency slots and an overlapping contention period where both types of MSs contend.

The results are shown in Figure 7.9 where we see further improvements over separate contention zones to 53.12% to 54.28%, 60.06% to 64.58% and 73.72% to 77.80% of emergency MSs processed for the overlapping contention windows with dedicated emergency contention zones of 25%, 50% and 75% respectively. The results are summarized in Table 7.3.

Percentage of Emergency MSs Processed	
WiMAX Default	49.44% ±0.18% to 49.61% ± 0.08%
25% Emergency CW	53.12% ±0.08% to 54.28% ± 0.11%
50% Emergency CW	60.06% ±0.25% to 64.58% ± 0.13%
75% Emergency CW	73.72% ±0.09% to 77.80% ± 0.19%

Table 7.3: Percentage of emergency MSs processed within 50% of all MSs - overlapping contention window.

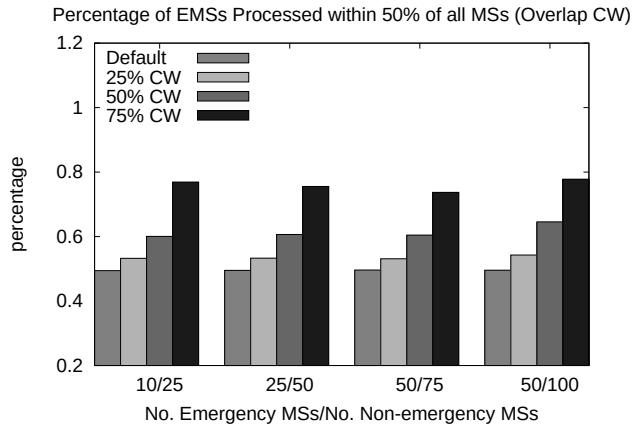


Figure 7.9: Percentage of emergency MSs processed in the first 50% of all MSs - overlapping contention window.

Both graphs in Figure 7.10 compare the percentage of emergency/non-emergency MS collisions between the WiMAX default setting versus when the BS sets the dedicated emergency contention zone to 25%, 50% and 75%, but allows emergency MSs to contend across all contention slots. The upper graph shows the simulation for a single frame of contention, where all MSs arrive at once. This shows a similar result to the expected as shown in Figure 7.6. The lower graph shows the same measurement with the standard multiple frame arrivals. In both cases, we can see that the probability of collisions between emergency and non-emergency MSs can be greatly reduced simply by reserving a portion of the contention period for emergency use.

7.5.2 Emergency CDMA Ranging Codes

In our simulations, the use of CDMA ranging codes does not have a major impact in the percentage of emergency MSs processed before non-emergency MSs. The exception is in extreme cases where with a large emergency contention window size, large number of emergency CDMA codes along with a large MSs arrivals lead to a greater than 30% collision rate among non-emergency MSs. The main use of emergency CDMA codes is that they are required in order to distinguish between emergency and non-emergency MSs when they have overlapping contention periods. One benefit of increasing the emergency CDMA codes is to reduce the number of collisions among

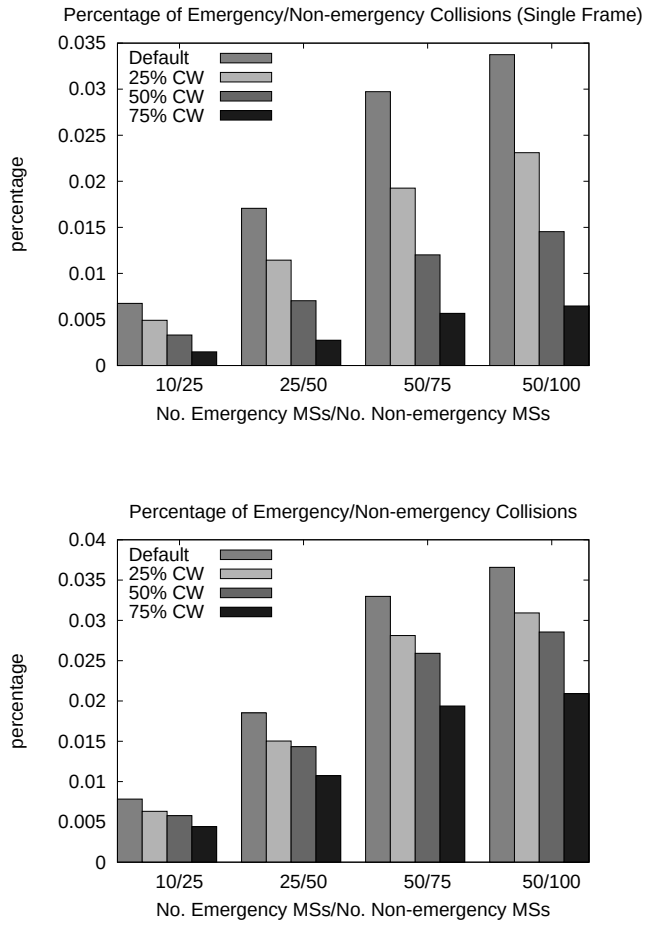


Figure 7.10: Probability of MS collisions, overlapping contention window.

emergency MSs as shown in Figure 7.11 for the cases of a 50% emergency contention zone with overlap, 25% and 50% assignment of the CDMA codes for emergency MSs.

7.5.3 Base Station Delayed Response

In our final set of experiments, we investigated the effect of introducing a BS delay when sending a response to non-emergency MSs. The delay is for a portion of the default ranging timeout, T3, while the BS is waiting to hear from possibly more emergency MSs. In Figure 7.12, we present the results for introducing a BS delay of 10, 20 and 30 ms with a 50% separate emergency contention zone. Here we see that a delay of only 20 ms increases the percentage of emergency MSs processed within 50%

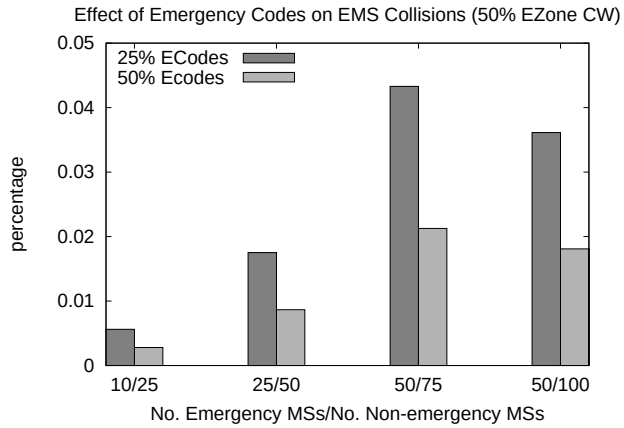


Figure 7.11: Effect of assigned emergency CDMA ranging codes.

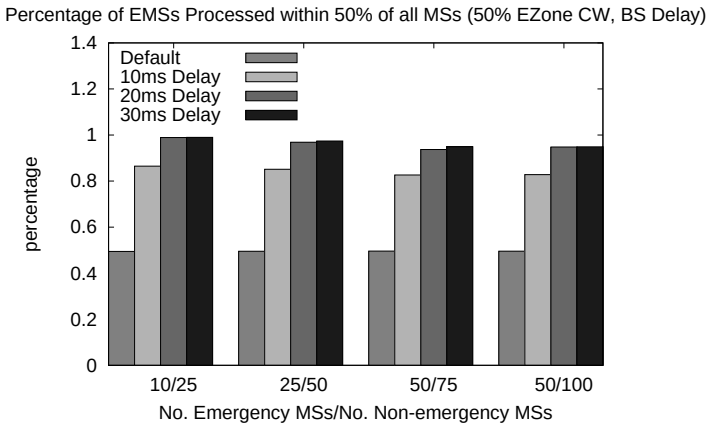


Figure 7.12: Percentage of emergency MSs processed in the first 50% of all MSs - split contention windows, BS delay.

of all MSs to the range of 93.70% to 98.91%. The results are further summarized in Table 7.4.

7.6 Conclusions and Future Work

We have proposed strategies that can enable BSs in a WiMAX/ 802.16 network to provide prioritized network entry access to an emergency class of MS. Our proposed strategies reduce or eliminate the interference between emergency and non-emergency

Percentage of Emergency MSs Processed	
WiMAX Default	49.44% ±0.18% to 49.61% ± 0.08%
10 ms Delay	82.64% ±0.13% to 86.49% ± 0.24%
20 ms Delay	93.70% ±0.12% to 98.91% ± 0.12%
30 ms Delay	94.86% ±0.12% to 98.99% ± 0.12%

Table 7.4: Percentage of emergency MSs processed within 50% of all MSs - BS delayed response with split 50% contention window.

MSs during the contention-based CDMA initial ranging process. They can also let the BSs determine the type of MS attempting initial ranging and gives the BSs flexibility on controlling which MSs are permitted to continue the network entry process even before the actual ID of the MSs are known.

With the introduction of emergency contention zones we increased the percentage of emergency MSs range requests processed by between 24% to 53% when compared to the WiMAX/802.16 default. These percentages are increased to between 70% and 94% over default WiMAX/802.16 with a 50% emergency contention zone and a BS delayed range response to non-emergency MSs of 10 ms and 20 ms respectively.

The implementation of our strategies is flexible, where BSs can operate from WiMAX default mode where all MSs have equal access through to total emergency mode where only emergency MSs attempt the network entry process. In times of high emergency demand, or disaster, our proposed strategies can be tuned to only grant access to emergency MSs. The strategies presented can also be applied to other contention regions such as those for periodic ranging and bandwidth requests.

Continuing work includes a more extensive simulation to provide for the on-the-fly tweaking of our proposed strategies to determine thresholds for the setting of emergency contention zone sizes, emergency CDMA codes assigned and timings of BS delays. This would allow the BSs to adjust to the real-time network conditions. Larger scale evaluation should be done including multiple channels per BS as well as having multiple collaborating BSs in order to determine how to best handle the arrival of an explosive number of emergency MSs as emergency situations arise. It should be investigated how other ranging types as well as bandwidth requests from emergency MSs can benefit from similar strategies. Additional future work includes investigating the the WiMAX/802.16 Quality of Service (QoS) structure in order to

better understand and support the QoS requirements of emergency applications as they compete with non-emergency applications with similar QoS demands. The work in this Chapter is to be submitted for publication [43].

Part IV

Conclusion

Chapter 8

Conclusions and Future Work

Our goal was to investigate problems of access in next-generation broadband wireless networks. We proposed solutions and evaluated the performance for the following important problem areas:

- Efficient Access to Data through Filtering
- Improved Network Access in Next-generation Broadband Wireless Networks
- Prioritized Emergency Network Access in Next-generation Broadband Wireless Networks

The following sections summarize our contributions and discuss some continuing work in the areas of study.

8.1 Contributions

8.1.1 Efficient Filtering of Data

In our work on improving access to data, we first compiled a comprehensive survey on forwarding engines for XML documents in publish/subscribe content-based routing systems [37]. We also have introduced our Hybrid XML filtering engine for content-based routing. Our filtering engine examines the type of XPath queries and stores queries of different types in different structures. We compared the performance of our filtering engine with that of AFilter [53] and observed improvements in all of our key measurements. XML document matching times for the Hybrid filtering engine showed between a 23% and 32% improvement over AFilter. The build time of the query structure has been reduced by between 13%-18% and the memory usage has seen similar reductions with our Hybrid filter over that of AFilter.

8.1.2 Scanning, Ranging and Handovers in WiMAX/802.16

We have examined the scanning and ranging operations for a WiMAX/802.16 MS. We have introduced two new strategies to aid in reducing the number of frequencies to check while scanning to find a downlink from a BS and provided some experimental results of a simulation based on real-world mobility traces. In the two scenarios we tested, *along the path* vs. *random locations*, we have shown that the *MRU* (*MFU* respectively) performed best when compared against the default WiMAX/802.16 scanning strategy. The proportion of time that scanning was completed within checking ten frequencies has been increased from 51.9% (45.6%) for the WiMAX default to 86.9% (70.7%) for the *MRU* strategy and 79.5% (79.5%) for the *MFU* strategy for along the path (random points) scenarios.

We introduced an additional strategy to enhance the neighbor advertisement messages sent by the BS to assist the MS with handovers. In this strategy, the MS builds a history of previous handovers and uses this information to select which of the advertised neighboring BSs is most likely the correct one for a handover. We have shown that the *previous handover* strategy improves performance significantly over that of the WiMAX/802.16 default strategy. On average, the *previous handover* strategy chooses the correct target BS on its first attempt 72.38% of the time as compared to only 22.78% of the time for the WiMAX/802.16 default.

Finally, we have introduced a series of initial ranging strategies to aid in reducing the number of ranging messages sent during the ranging process. In the two scenarios we tested, *along the path* vs. *random locations*, we have shown that the *MRU* (*MFU* respectively) performed best when compared against the default and exponential increment ranging strategies. The average number of ranging messages required has been reduced from 7.28 (7.29) for the WiMAX default to 1.20 (1.55) for the *MRU* strategy and 1.35 (1.36) for the *MFU* strategy for along the path (random points). Additionally, with by using the *MRU* strategy for ranging, the power expended has been reduced by between 70.9% and 73.3% over the default WiMAX strategy.

Our work is complementary to the earlier research presented such as that done by [58, 78, 83, 90, 104]. Our strategies are MS based and require no additional network support. This initial work on scanning, ranging and handovers was the basis for

our work on developing our TOD and LPT mobility profiles. This work has been published in [40] and [41].

8.1.3 Mobility Profiles for Scanning and Handovers

We have focused on improving the scanning times for WiMAX/802.16 MSs during handovers. We have introduced two mobility profiles, *time-of-day* and *location plus trajectory*, using our MFU scanning strategy, to aid in reducing the number of frequencies checked while scanning to find a downlink from a BS. We have provided some performance results from a set of experiments based on simulated and real-world mobility traces.

Through the simulated mobility experiments, we have shown that using either TOD or LPT mobility profiles improves handover scanning performance over that of the WiMAX/802.16 default and our previous work. For example, the worker class using the Manhattan mobility model could complete scanning within five frequencies on average 50.29%, 73.23% and 78.71% of the time for the TOD, LPT-16 and LPT-25 mobility profiles respectively. This is an increase over the 43.73% seen from our original MFU and the 28.96% for the WiMAX default. Additionally, the LPT mobility profile provided significant improvement for the Wanderer mobility class where the TOD profile did not. We observed MSs completing scanning within five frequencies 47.65% and 54.58% of the time on average for LPT-16 and LPT-25 as compared to the 26.68% seen for our original MFU strategy.

Using the real-world APRS mobility data, along with the LPT mobility profiles, the MS can complete scanning within five frequencies 55.53% and 63.99% of the time for LPT-16 and LPT-25 respectively. This is an improvement over the 32.43% for our original MFU strategy and the 23.36% for the WiMAX default strategy.

We also show improvement in scanning performance when using the TOD mobility profile for the *previous handover* strategy when simulating with the APRS mobility trace data. However, a more sophisticated set of TOD mobility profiles should be implemented. In addition, other sets of mobility data should be examined, such as users of cellular telephones. This work is to be submitted for publication [42].

8.1.4 Prioritized Emergency Access

We have proposed tunable strategies that can enable BSs in a WiMAX/802.16 network to provide prioritized network entry access to an emergency class of MS. Our proposed strategies reduce or eliminate the interference between emergency and non-emergency MSs during the anonymous contention-based CDMA initial ranging process. They also let the BS determine the type of MS attempting initial ranging and gives the BS flexibility on controlling which MSs are permitted to continue the network entry process even before the actual ID of the MSs are known.

With the introduction of emergency contention zones we increased the percentage of emergency MSs range requests processed by between 24% to 53% when compared to the WiMAX/802.16 default. These percentages are increased to between 70% and 94% over default WiMAX/802.16 with a 50% emergency contention zone and a BS delayed range response to non-emergency MSs of 10ms and 20ms respectively.

The implementation of our strategies is flexible. BSs can operate from a WiMAX default mode where all MSs have equal access through to a total emergency mode where only emergency MSs will attempt the network entry process. In times of high emergency demand, or disaster, our proposed strategies can be tuned to only grant access to emergency MSs. The strategies presented can also be applied to other contention regions such as those for periodic ranging and bandwidth requests. This work is to be submitted for publication [43].

8.2 Future Work

8.2.1 Efficient Filtering of Data

There are still a number of open problems that remain in the area of XML filtering. Implementation of prefix-caching and suffix-clustering [53] would further improve matching times. Some empirical studies on the actual "query mix" for XPath queries for publish/subscribe systems should be performed in order to better understand their nature.

There may be other strategies that help in improving filtering performance. Examining the possibility of decomposing the complex queries further into ones with

wild cards, ones with descendent operators or ones with both. Alternatively, we can examine the effect of storing XPath queries based on their depth.

The implementation of the AFilter filtering engine used for comparison as well as to store complex queries in our Hybrid filtering engine was not optimized and thus had less than the expected performance. Finally, the main focus of most previous and current work is on *structure* matching. The optimization of *value* matching needs to be investigated in the context of an XML matching engine. There may be many queries that have overlapping prefixes, resulting in many different values that must be compared at each common location step.

8.2.2 Mobility Profiles using Scanning and Handovers

Future work includes providing a larger set of finer grained time-of-day mobility profiles. There is the additional problem of developing an MS algorithm for building its optimal set of time-of-day profiles. The set of time periods that need to be profiled depends on the mobility patterns of individual MSs.

A further extension to the LPT model is to include building profiles for different routes an MS may take. The MS could then determine which route it is currently following and use this on which to base its scanning decisions. Additionally, the assignment of LPT zones to the coverage area could be done differently by examining the MSs routes and determining locations of roads etc.

Finally, an additional refinement to the model to investigate how different loads on the WiMAX/802.16 BSs, including multiple MSs attempting to perform ranging with the BS during the same ranging opportunities, may affect the scanning and ranging times.

8.2.3 Prioritized Emergency Access

Continuing work towards improving network access for emergency MSs includes a more extensive simulation to provide for the on-the-fly tweaking of our proposed strategies to determine thresholds for the setting of emergency contention zone sizes, emergency CDMA codes assigned and timings of BS delays. This would allow the BSs to adjust to the real-time network conditions. Larger scale evaluation should be done

including multiple frequencies per BS as well as having multiple collaborating BSs in order to determine how to best handle the arrival of an explosive number of emergency MSs as emergency situations arise. It should be investigated how other ranging types as well as bandwidth requests from emergency MSs can benefit from similar strategies. Additional future work includes investigating the the WiMAX/802.16 Quality of Service (QoS) structure in order to better understand and support the QoS requirements of emergency applications as they compete with non-emergency applications with similar QoS demands.

Bibliography

[1] 3GPP. The 3rd Generation Partnership Project, 3GPP Work Item - Priority Multimedia Service.
http://www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGS_19/Docs/PDF/SP-030039.pdf.

[2] 3GPP. The 3rd Generation Partnership Project, 3GPP Work Item - PS domain and IMS impacts for supporting IMS emergency calls.
http://www.3gpp.org/specs/WorkItem-info/WI–32045.htm.

[3] 3GPP. 3GPP TR 22.950 The 3rd Generation Partnership Project, 3GPP Work Item - Priority Service Feasibility Study.
http://www.3gpp.org/ftp/Specs/html-info/22950.htm, 2007.

[4] AVAYA Mobile Communication System.
http://www.avaya.com, 2006.

[5] Basex Inc., "Information Overload Strategies: The Impact On The Knowledge Organization," *Technical Report*, 2007.

[6] ETSI. EMTEL - Emergency Telecommunications.
http://www.emtel.etsi.org/.

[7] ETSI. ETSI TS 102 424 (2005-09) - "Requirements for the NGN network to support emergency communication from citizen to authority," *http://webapp.etsi.org/WorkProgram/Report_WorkItem.asp?WKI_ID=20616*, 2005.

[8] IEEE 802.16-2008, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems," *IEEE Std 802.16-2008*, 2008.

[9] IEEE 802.16-2005, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems - Amendment 2:Physical and medium access control layers for combined fixed and mobile operation in licensed bands annd corrigendum 1",. *IEEE Std 802.16-2005*, 2005.

[10] IEEE 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems," *IEEE Std 802.16-2004 (Revision of IEEE Std 802.16-2001)*, 2004.

[11] IEEE 802.16j, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems - Multihop Relay Specification," unapproved draft document, *IEEE Std 802.16j*, 2007.

[12] IEEE 802.11p - Draft Amendment to Standard for Information technology - Telecommunications and information exchange between systems - LAN/MAN Specific Requirements - Wireless Access in Vehicular Environments (WAVE). 2005.

[13] IEEE P1609.3 Standard for Wireless Access in Vehicular Environments (WAVE) - Networking Services. *IEEE Std 1609.3*, pages 1–93, 2005.

[14] IETF. Emergency Context Resolution with Internet Technologies (ECRIT), IETF Working Group. *http://www.ietf.org/dyn/wg/charter/ecrit-charter.html*.

[15] ITU. International Telecommunication Union. *http://www.itu.int*.

[16] LSI. LSI Tarari Content Processor. *http://www.lsi.com/networking_home/networking_products/ tarari_content_processors/xml/index.html*

[17] MOBIDIS - Universita' Degli Studi Di Roma "La Sapienza", *Dipartimento Di Informatica e Sistemistica. http://www.dis.uniroma1.it/pub/mecella/projects/MobiDIS/index.html*.

[18] NIST, Distributed Testbed for First Responders. *http://w3.antd.nist.gov/first_responders_news.shtml*.

[19] Nyquetek Inc., "Wireless Priority Service for National Security/Emergency Preparedness: Algorithms for Public Use Reservation and Network Performance," 2002. *http://wireless.fcc.gov/releases/da051650PublicUse.pdf.*

[20] Terrestrial Trunked Radio - TETRA. *http://www.tetra-association.com/.*

[21] WIDENS - WIreless DEployable Network System. *http://www.comlab.hut.fi/projects/WIDENS/.*

[22] XML Path Language (XPath) Version 1.0. *http://www.w3.org/TR/xpath*.

[23] K. A. Agha, P. Jacquet, and N. Vvendenskaya. "Analysis of the Priority Stack Random-Access Protocol in W-CDMA Systems," *IEEE Transactions on Vehicular Technology*, 51(3):588–596, 2002.

[24] M. Aguado, E. Jacob, P. Saiz, J. Matias, M. Higuero, N. Toledo and M. Berbineau, "Scanning on handover enhancement issues in video application deployments on WiMAX Mobile Networks", *In Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, pages 1–7, 2008.

[25] D. Agrawal and Q.-A. Zeng. *Introduction to Wireless and Mobile Systems*. Thomson-Engineering, first edition, 2002.

[26] M. K. Aguilera, R. E. Strom, D. C. Sturman, M. Astley, and T. D. Chandra. "Matching events in a content-based subscription system," *In Proceedings of the Symposium on Principles of Distributed Computing*, pages 53–61, 1999.

[27] M. Altinel and M. J. Franklin. "Efficient filtering of XML documents for selective dissemination of information," *In Proceedings of the 26th International Conference on Very Large Data Bases (VLDB)*, pages 53–64, 2000.

[28] F. D. Angelis, I. Habib, F. Davide, and M. Naghshineh. "Increasing revenues via adaptive filtering in wired/wireless networks," *Computer Networks*, 50(9):1261–1280, 2006.

[29] G. Ashayer, H. K. Y. Leung, and H.-A. Jacobsen. "Predicate matching and subscription matching in publish/subscribe systems," *In Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops*, pages 539–546, 2002.

[30] S. Bangolae, C. Bell, and E. Qi. "Performance Study of Fast BSS Transition using IEEE 802.11r," *In Proceedings of the International Conference On Communication And Mobile Computing,* pages 737–742, 2006.

[31] M. Barbeau, P. Boone, and E. Kranakis. "WiMAX/802.16 Broadband Wireless Networks," *Selected Topics in Communication Networks and Distributed Systems,* World Scientific, Ed. S. Misra, S. C. Misra and I. Woungang, first edition, 2010.

[32] C. Beard. "Dynamic Agent-Based Prioritized Resource Allocation for Stressed Networks," *Ph.D. Dissertation*, 1999.

[33] C. Beard and V. Frost. "Prioritization of Emergency Network Traffic using Ticket Servers: A Performance Analysis," *Simulation: Transactions of the Society for Modeling and Simulation International*, 80(6), pages 289–299, 2004.

[34] C. Beard, "Preemptive and Delay-Based Mechanisms to Provide Preference to Emergency Traffic." *Computer Networks Journal*, Vol. 47:6, pages 801–824, 2005.

[35] Z. Becvar, P. Mach and R. Bestak, "Optimization of handover scanning procedure in WiMAX networks with relay stations," *In Proceedings of the 3rd International Symposium on Wireless Pervasive Computing,* pages 581–585, 2008.

[36] S. Bittner and A. Hinze. "On the benefits of non-canonical filtering in publish/subscribe systems," *In Proceedings of the Fourth International Workshop on Distributed Event-Based Systems (DEBS) (ICDCSW'05)*, pages 451–457, 2005.

[37] P. Boone. "Survey of Filtering Engine Techniques for Publish/Subscribe Systems," *Internal Alcatel-Lucent Document*, pages 1–36, 2006.

[38] P. Boone. "XML Filtering Engine Structure Proposal," *Internal Alcatel-Lucent Document*, pages 1–11, 2006.

[39] P. Boone. "A Hybrid XML Filtering Engine for Publish/Subscribe Content-Based Routing," *Internal Alcatel-Lucent Document*, pages 1–26, 2007.

[40] P. Boone, M. Barbeau, and E. Kranakis, "Strategies for Fast Scanning and Handovers in WiMAX/802.16," *In Proceedings of the Second International Conference on Networks and Workshops,* AccessNets 2007, pages 1–7, 2007.

[41] P. Boone, M. Barbeau, and E. Kranakis, "Strategies for Fast Scanning, Ranging and Handovers in WiMAX/802.16," *International Journal of Communication Networks and Distributed Systems,* Vol. 1, No.4/5/6, pages 414–432, 2008.

[42] P. Boone, M. Barbeau, and E. Kranakis, "Using Time-of-Day and Location-based Mobility Profiles to Improve Scanning During Handovers in WiMAX," submitted.

[43] P. Boone, M. Barbeau, and E. Kranakis, "Prioritized Access for Emergency Stations in Next Generation Broadband Wireless Networks," submitted.

[44] B. Bou-Diab and P. Boone. Method and Appliance for XML Policy Matching. *USPTO Application number: 20090037379. Filed July 2007.*

[45] A. Boukerche, S. K. Das, and A. Fabbri. "SWiMNet: A Scalable Parallel Simulation Testbed for Wireless and Mobile Networks," *Wireless Networks*, 7:467–486, 2001.

[46] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau. Extensible Markup Language (XML) 1.1 (Second Edition). *http://www.w3.org/TR/xml11/.*

[47] Gene Brewer, Automatic Position Reporting System: A Basic Overview, *http://www.w4aaz.org/attachments/APRS_Overview_rev2.pdf*

[48] Bob Bruninga, Automatic Packet/Position Reporting System (APRS), *http://www.usna.edu/Users/aero/bruninga/APRS-docs/APRS.TXT*, 2002.

[49] B. Bruninga. "A Description of APRS by WB4APR, and more!," Packet Users Notebook,. *CQ Magazine*, 1993.

[50] B. Bruninga, "Automatic AX.25 Position and Status Reporting," *In Proceedings of the 11th Computer Networking Conference,* 1992.

[51] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli. "Integration of 802.11 and Third-Generation Wireless Data Networks," *In Proceedings of the Twenty-second AnnualJoint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pages 503–512, 2003.

[52] A. Campailla, S. Chaki, E. M. Clarke, S. Jha, and H. Veith. "Efficient Filtering in Publish-Subscribe Systems Using Binary Decision Diagrams," *In Proceedings of the International Conference on Software Engineering*, pages 443–452, 2001.

[53] K. S. Candan, W.-P. Hsiung, S. Chen, J. Tatemura, and D. Agrawal. "AFilter: Adaptable XML filtering with Prefix-Caching and Suffix-Clustering," *In Proceedings of the 32nd international conference on Very large data bases (VLDB)*, pages 559–570, 2006.

[54] A. Carzaniga, M. Rutherford, and A. L. Wolf. "A Routing Scheme for Content-based Networking," *In Proceedings of the Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2:918–928, 2004.

[55] A. Carzaniga and A. L. Wolf. "Forwarding in a Content-based Network," In *In Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications (SIGCOMM)*, pages 163–174, 2003.

[56] C.-Y. Chan, P. Felber, M. Garofalakis, and R. Rastogi. "Efficient Filtering of XML Documents with XPath Expressions," *In Proceedings of the 18th International Conference on Data Engineering (ICDE)*, pages 235–244, 2002.

[57] H-T. Chou and D. J. Dewitt. "An evaluation of buffer management strategies for relational database systems," *In Proceedings of the 11th international conference on Very Large Data bases (VLDB)*, pages 127–141, 1985.

[58] S. Choi, G. Hwang, T. Kwon, A. Lim, and D. Cho. "Fast Handover Scheme for Real-time Downlink Services in IEEE 802.16e BWA systems," *In Proceedings of the Vehicular Technology Conference (VTC)*, pages 2028–2032, 2005.

[59] M. . Chuah, Q. Zhang, and O. Yue. "Access priority schemes in UMTS MAC," *In Proceedings of the Wireless Communications and Networking Conference (WCNC),* Vol.2, pages 781–786, 1999.

[60] M. D. Corner and B. D. Noble. "Zero-interaction authentication," *In Proceedings of the 8th annual international conference on Mobile computing and networking (MobiCom)* , pages 1–11, 2002.

[61] F. Dellutri, G. Me, and M. Strangio. "Local Authentication with Bluetooth enabled Mobile Devices," *In Proceedings of the Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services (ICAS-ICNS)* , pages 72–72, 2005.

[62] Y. Diao, M. Altinel, M. J. Franklin, H. Zhang, and P. Fischer. "Path Sharing and Predicate Evaluation for High-performance XML Filtering," *ACM Transactions on Database Systems*, 28(4):467–516, 2003.

[63] W. Diffie and M. Hellman. "New directions in cryptography," *IEEE Transactions on Information Theory 22*, pages 644–654, 1976.

[64] C. Eklund, R. B. Marks, K. L. Stanwood, and S. Wang. "IEEE Standard 802.16: A Technical Overview of the Wirelessman™Air Interface for Broadband Wireless Access," *IEEE Communications Magazine*, pages 98–107, 2002.

[65] F. Fabret, H. A. Jacobsen, F. Llirbat, J. Pereira, K. A. Ross, and D. Shasha. "Filtering Algorithms and Implementation for Very Fast Publish/Subscribe Systems," *In Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data*, pages 115–126, 2001.

[66] F. Fabret, F. Llirbat, J. Pereira, and D. Shasha. "Efficient Matching for Content-based Publish/Subscribe Systems," Technical Report, INRIA. http://wwwcaravel.inria.fr/pereira/matching.ps.

[67] M. F. Fernandez and D. Suciu. "Optimizing Regular Path Expressions Using Graph Schemas," *In Proceedings of the International Conference on Data Engineering (ICDE)*, pages 14–23, 1998.

[68] R. Filjar and S. Desic. "Architecture of the Automatic Position Reporting System (APRS)," *In Proceedings of the 46th International Symposium Electronics in Marine (ELMAR)*, pages 331–335, 2004.

[69] C. Fleischer.
"The History of Police Communications,"
http://www.ci.irvine.ca.us/ipd/divisions/dispatch/911_dispatch.asp.

[70] T. Fujiwara, N. Iida, and T. Watanabe. "An Ad-hoc Routing Protocol in Hybrid Wireless Networks for Emergency Communications," *In Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 748–754, 2004.

[71] S. Ganguly, S. Bhatnagar, A. Saxena, R. Izmailov, and S. Banerjee. "A Fast Content-based Data Distribution Infrastructure," *In Proceedings of the 25th IEEE International Conference on Computer Communications (INFOCOM)*, pages 1–13, 2006.

[72] L. M. Gavrilovska and V. M. Atanasovski. "Enabling Communications in Emergency Scenarios," *In Proceedings of the 8th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services (TELSIKS)*, pages 67–76, 2007.

[73] X. Gong, W. Qian, Y. Yan, and A. Zhou. "Bloom Filter-Based XML Packets Filtering for Millions of Path Queries," *In Proceedings of the 21st International Conference on Data Engineering (ICDE)*, pages 890–901, 2005.

[74] T. J. Green, A. Gupta, G. Miklau, M. Onizuka, and D. Suciu. "Processing XML Streams with Deterministic Automata and Stream Indexes," *ACM Transactions on Database Systems*, 29(4):752–788, 2004.

[75] T. A. Horan and B. L. Schooley. "Time-Critical Information Services," *Communications of the ACM*, 50(3):73–78, 2007.

[76] P. Hosein, "Dynamic resource allocation for uplink contention channels in WiMAX," *In Proceedings of the Wireless Telecommunications Symposium*, pages 25-30, 2008.

[77] S. Hou and H.-A. Jacobsen. "Predicate-based Filtering of XPath Expressions," *In Proceedings of the 22nd International Conference on Data Engineering (ICDE'06)*, pages 53-62, 2006.

[78] F.-Y. Hsieh, Y.-W. Chen, and P-W.Wu. "Cross Layer Design of Handoffs in IEEE 802.16e Networks," *In Proceedings of the International Computer Symposium (ICS)*, pages 703–708, 2006.

[79] M. Inoue, G. Wu, K. Mahmud, H. Murakami, and M. Hasegawa. "Development of MIRAI system for Heterogeneous Wireless Networks," *In Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, pages 69–73, 2002.

[80] H. Jang, J. Jee, Y.-H. Han, S. Park, and J.Cha. "Mobile IPV6 Fast Handovers over IEEE 802.16e Networks," *MIPSHOP Working Group, Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-mipshop-fh80216e-01.txt*, 2007.

[81] D. B. Johnson and D. A. Maltz. "Dynamic Source Routing in Ad Hoc Wireless Networks," *Mobile Computing*, 353:153–181, 1996.

[82] H.-D. Kim, B.-S. Bae, H.-H. Choi, and D.-H. Cho. "A New CPCH Access Scheme for Priority Service," *IEICE Transactions on Communications*, E86-B(4):1448–1452.

[83] K. Kim, C. Kim, and T. Kim. "A Seamless Handover Mechanism for IEEE 802.16e Broadband Wireless Access," *In Proceedings of the International Conference of Computational Science (ICCS)*, pages 527-534, 2005.

[84] K. Kutzner, J.-J. Tschouto, M. Bechler, L. Wolf, and B. Bochow. "Connecting Vehicle Scatternets by Internet-Connected Gateways," *In Proceedings of the Workshop on Multiradio Multimedia Communications (MMC)*, 2003.

[85] J. Kwon, P. Rao, B. Moon, and S. Lee. "FiST: Scalable XML Document Filtering by Sequencing Twig Patterns," *In Proceedings of the 31st international conference on Very Large Data Bases*, pages 217–228, 2005.

[86] L.-T. Lee and C.-F. Wu. "An HMM-based call admission control policy for supporting QoS in wireless cellular networks," *Computers and Electrical Engineering*, 33(3):153–165, 2007.

[87] B. Lin, P.-H. Ho, L.-L. Xie, and X. Shen. "Optimal Relay Station Placement in IEEE 802.16j Networks," *In Proceedings of the International Wireless Communications and Mobile Computing Conference, IWCMC '07*, pages 25–30, 2007.

[88] C. Lindemann, M. Lohmann, and A. Thummler. "Adaptive Call Admission Control for QoS/Revenue Optimization in CDMA Cellular Networks," *Wireless Networks*, 10:457–472, 2004.

[89] M. Lott, R. Halfmann, E. Schulz, and M. Radimirsch. "Medium Access and Radio Resource Management for Ad hoc Networks Based on UTRA TDD," *In Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC)*, pages 76–86, 2001.

[90] H. Mahmoud, H. Arslan, and M. Ozdemir. "Initial Ranging for WiMAX (802.16e) OFDMA," *In Proceedings of 25th Annual Military Communications Conference, MILCOM,* pages 1–7, 2006.

[91] H. Mahmoud, H. Arslan, and M. Ozdemir. "An Efficient Initial Ranging for WiMAX (802.16e) OFDMA," *Computer Communications 32*, pages 159-168, 2009

[92] B. Manoj and A. Baker. "Communication Challenges in Emergency Response," *Communications of the ACM*, 50(3):51–53, 2007.

[93] S. McCann. "IEEE 802 Emergency Services (ES) Call for Interest (CFI,)" *http://www.ieee802.org/802_tutorials/07-November/IEEE 802 Emergency Services CFI.ppt*, 2007

[94] M. Mishra and A. Sahoo. "A Contention Window Based Differentiation Mechanism for providing QoS in Wireless LANs," *In Proceedings of the International Conference on Information Technology (ICIT)*, pages 72-76, 2006.

[95] I. Murase, M. Murano, and H. Ohno. "Standardization Activity on Emergency Telecommunication Systems in ITU-T," *In Proceedings of the 2004 International Symposium on Applications and the Internet Workshops*, pages 217–222, 2004.

[96] P. Neves, P. Simoes, A. Gomes, L. Mario, S. Sargento, F. Fontes, E. Monteiro, and T. Bohnert. "WiMAX for Emergency Services: An Emperical Evaluation,"

In Proceedings of the 2007 International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST), pages 340–345, 2007.

[97] N. Mekhiel, "Multi-level cache with most frequently used policy: A new concept in cache design" *In Proceedings of the International Conference on Computer Applications in Industry and Engineering,* 1995.

[98] D. Niyato and E. Hossain. "QoS-aware bandwidth allocation and admission control in IEEE 802.16 broadband wireless access networks: A non-cooperative game theoretic approach," *Computer Networks*, 51:3305–3321, 2007.

[99] B. Noble and M. Corner. "The case for transient authentication," *In Proceedings of the 10th ACM SIGOPS European Workshop,* pages 24–29, 2002.

[100] J. M. Peha. "How America's Fragmented Approach to Public Safety Wastes Money and Spectrum," *Telecommunications Policy*, 31(10-11):605–618, 2007.

[101] J. Peng and L. Cheng. "Priority-ensured Medium Access in Ad Hoc Networks," *In Proceedings of the IEEE International Performance, Computing, and Communications Conference, IPCCC 2007.*, pages 170–177, 2007.

[102] P. Rao and B. Moon. "Prix: Indexing and Querying XML Using Prufer Sequences," *In Proceedings of the 20th International Conference on Data Engineering*, page 288, 2004.

[103] C. Redding and C. Taylor. "Priority access service in cellular and PCS networks," *In Proceedings of 20th Annual Military Communications Conference (MILCOM), Communications for Network-Centric Operations: Creating the Information Force*, 1:586–590 vol.1, 2001.

[104] R. Rouil and N. Golmie. "Adaptive Channel Scanning for IEEE 802.16e," *In Proceedings of 25th Annual Military Communications Conference (MILCOM)*, pages 1–6, 2006.

[105] D. Shi and C. Tang. "A Fast Handoff Scheme Based on Local Authentication In Mobile Network," *In Proceedings of the 6th International Conference on ITS Telecommunications Proceedings*, pages 1025–1028, 2006.

[106] D. Shrestha, S.-H. Lee, S.-C. Kim, and Y.-B. Ko. "New Approaches for Relay Selection in IEEE 802.16 Mobile Multi-hop Relay Networks," *In Proceedings of Euro-Par 2007 - Parallel Processing*, pages 950–959, 2007.

[107] Z. Tao, K. Teo, and J. Zhang. "Aggregation and Concatenation in IEEE 802.16j Mobile Multihop Relay (MMR) Networks," *In Proceedings of the IEEE Mobile WiMAX Symposium*, pages 85–90, 2007.

[108] J.M. Thorington Jr. and D. J. Irwin, "An Adaptive Replacement Algorithm for Paged Memory Computer Systems," *In Proceedings of the IEEE Transactions on Computers*, Vol. C-21, No. 10, pages 1053–1061, October 1972.

[109] P. van de Berg. Carrier scanning in cellular networks. *Patent number: WO/2000/078075, World Intellectual Property Organization, http://www.wipo.int*, 2000.

[110] H. Wu, C. Qiao, S. De, and O. Tonguz. "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE Journal on Selected Areas in Communications*, 19(10):2105–2115, 2001.

[111] T. W. Yan and H. García-Molina. "Index Structures for Selective Dissemination of Information Under the Boolean Model," *In Proceedings of the ACM Transactions on Database Systems*, 19(2):332–364, 1994.

[112] K. K. A. Zahid, L. Jun, K. Kazaura, and M. Matsumoto. "IP Network for Emergency Service," *In Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia*, pages 165–170, 2004.

[113] J. Zeng; H. Minn, "An Investigation into Initial Ranging Method for Mobile OFDMA Systems," *In Proceedings of the IEEE Sarnoff Symposium*, pages 1-5, 2008.

[114] Y. Zhang and D. Liu. "An adaptive algorithm for call admission control in wireless networks," *In Proceedings of the IEEE Global Telecommunications Conference*, Vol. 6, pages 3628–3632, 2001.

[115] B. Zhou, K. Xu, M. Gerla, "Group and swarm mobility models for ad hoc network scenarios using virtual tracks," *In Proceedings of the IEEE Military Communications Conference,* Vol. 1, pages 289–294, 2004.

[116] J. Zhou and C. Beard. "Tunable Preemption Controls for a Cellular Emergency Network," *In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC),* pages 3647–3652, 2007.

[118] J. Zhou and C. Beard, "Comparison of Combined Preemption and Queuing Schemes for Admission Control in a Cellular Emergency Network," *In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 122–128, 2006.

[118] J.Zhou, Y. Yang, D. Pang, X. Jin, J. Shi and Z. Li, "Contention Region Allocation Optimization in IEEE 802.16 OFDMA Systems," *In Proceedings of the The 10th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 197–200, 2007.

[119] Y. Zhou Z. Zhang X. Zhou, "OFDMA Initial Ranging for IEEE 802.16e Based on Time-domain and Frequency-domain Approaches," *In Proceedings of the International Conference on Communication Technology*, pages 1–5, 2006.

[120] X. Fu and H. Minn, "Initial Uplink Synchronization and Power Control (ranging process) for OFDMA systems," *In Proceedings of the IEEE Global Telecommunications Conference*, GLOBECOM, Vol. 6, pages 3999–4003, 2004.

[121] X. Fu, Y, Li and H. Minn, "A new ranging method for OFDMA systems," *In Proceedings of the IEEE Transactions on Wireless Communications*, Vol. 6, No. 2, pages 629–639, 2007.

[122] L. Sanguinetti, M. Morelli and H. V.Poor, "An Improved Scheme for Initial Ranging in OFDMA-based Networks," *In Proceedings of the 2008 IEEE International Conference on Communications*, pages 3469–3474, 2008.