

QoS-based Handover for Next Generation Wireless Networks

Sheetal Jadhav

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin,
New Zealand.

March 2013

Abstract

The deployment of the Next Generation Wireless Network (NGWN) involves different service providers, different radio access technologies and multi-mode mobile terminals that have to be compatible with existing services and technologies. It has provided many challenges for researchers and service providers. In particular, it is a difficult task to provide desired services, such as video streaming, teleconferencing, and data download/upload, with an acceptable Quality of Service (QoS) anywhere and any-time to the mobile users. These diverse needs of NGWN demand efficient and reliable technologies to satisfy users as well as network providers. Also NGWNs are expected to provide a high data rate and optimized QoS to multimedia and real-time applications over the Internet Protocol (IP) networks. However, due to the movement of the mobile terminals, seamless connectivity needs to be maintained when a mobile terminal moves across different cells or networks. Handover, which is the process of transferring an ongoing call from one base station to another, plays a critical role in achieving the above goals.

This thesis focuses on providing an end-to-end QoS to the users during the handover process in NGWNs. First, the thesis evaluates the performance of two kinds of popular wireless networks operating according to the WiMAX and UMTS standards, in terms of supporting Voice over Internet Protocol (VoIP) traffic. It then proposes a novel handover scheme compliant with the IEEE 802.21 standard (i.e. the Media Independent Handover protocol) which enables handover in an integrated network with UMTS and WiMAX. It takes into account the quality of a call and the load of the call among all the available access points while transferring the call between cells and networks. Mean Opinion Score (MOS) is used as the major metric of the call

quality in handover optimization. Comparing the novel MOS- based handover scheme with the traditional handover scheme based on Radio Signal Strength (RSS) through simulation of an integrated network of WiMAX and UMTS, the simulation results demonstrate that the proposed MOS-based scheme can maintain high call quality and reduce the probabilities of handover dropping and call dropping. Finally, the thesis proposes an energy efficient handover scheme that does not require frequent scanning of the network during handover process. Two schemes, heuristic and optimal, are proposed to select the optimal base station during handover. The simulation results show the handover scheme can significantly reduce energy consumption of mobile terminals.

Acknowledgements

Completion of my research work would not have been possible without the support of several people who supported me in one way or other. I express my sincere appreciation to you all, for your help and guidance during this roller coaster journey.

First of all, I am extremely grateful to my primary supervisor Dr. Zhiyi Huang for all his support and valuable guidance. His deep insights helped me during my research. During the initial stage of my research Zhiyi always made time for research meetings despite his busy schedule. It was a great opportunity for me to undertake my PhD under his guidance and it was a real pleasure working under him, with a lot of fun and excitement. I thank him for the funding support provided for my travel to the conferences.

I also thank sincerely my co-supervisor Dr. Haibo Zhang for his invaluable support in my research. He always made time available for clarifying the doubts and discussions especially during the conceptual stage of the research, and taught me how to write good research papers. I consider it was great opportunity to work with him and gain research expertise. His research advices, discussions and feedbacks have helped to strengthen my research.

I am very much indebted to my husband Bhaskar Ashoka. His invaluable support and love in every possible way helped me to complete my research work successfully. He supported me in all possible ways and gave me strength to complete this work. I thank my parents who always supported me to take up research studies and encouraged me at every stage of my academic and personal life.

I would like to thank the University of Otago and the Department of Computer Science for all the services provided to carry out my research. I would like to thank the research committee and the University Library for providing me with the required resources. I would like to thank my friends and colleagues, especially Jason Mair, Kai-Cheung , Tim, Jay, Syed Faisal and Jeremy for their insightful discussions and warm friendship.

Last but not least, I thank God for giving me strength, health, power and knowledge to take up and complete this research work.

Contents

1	Introduction	1
1.1	Preliminaries	2
1.2	Challenges in the Next Generation Wireless Networks	4
1.3	Motivation	7
1.4	Contributions	10
1.5	Thesis Outline	12
1.6	List of Publications	13
2	Background	14
2.1	NGWN Access Technologies	15
2.1.1	UMTS	15
2.1.2	WiMAX	19
2.1.3	LTE	22
2.2	Integrated Architectures	24
2.2.1	Tightly Coupled Architecture	25
2.2.2	Loosely Coupled Architecture	26
2.3	Handover in NGWN	27
2.3.1	Handover Overview	27
2.3.2	Handover Classification	28
2.3.3	Handover Process	29
2.4	Related Work	30
3	Performance Evaluation of Quality of VoIP in WiMAX and UMTS	33
3.1	Preliminaries	35
3.1.1	VoIP	35
3.1.2	QoS in UMTS	38
3.1.3	QoS in WiMAX	39
3.2	Simulation Setup	42
3.2.1	WiMAX Simulation Module	43
3.2.2	UMTS Simulation Module	44
3.2.3	Simulation Configuration	46
3.2.4	Performance Metrics	46
3.3	Simulation Results and Analysis	51
3.3.1	MOS	51
3.3.2	Packet End-to-End Delay	53
3.3.3	Jitter	55

3.3.4	Packet Delay Variation	58
3.4	Related Work	59
3.5	DISCUSSION	62
4	MOS-based Handover Protocol for Next Generation Wireless Networks	64
4.1	Related Work	67
4.2	E-model	70
4.2.1	Effect of Delay in Packet Network	70
4.2.2	Effect of Packet Loss Rate in Packet Network	72
4.3	Problem Formulation	72
4.4	Optimal Base Station Selection	75
4.5	Handover Protocol Design	78
4.5.1	IEEE 802.21 Standard	78
4.5.2	Parameter Acquisition	79
4.5.3	MOS-based Handover Protocol	80
4.6	Simulation Setup	84
4.7	Simulation Result	86
4.7.1	Mean Opinion Score (MOS)	86
4.7.2	Handover Dropping Probability (HDP)	88
4.7.3	Call Dropping Probability (CDP)	89
4.8	DISCUSSION	91
5	Energy Efficient Handover for NGWN	93
5.1	Introduction	93
5.2	Related Work	96
5.3	System Model	98
5.3.1	Energy on Data Communication	98
5.3.2	Energy on Channel Scanning	101
5.4	Problem Formulation	102
5.5	Proposed Solution	103
5.5.1	Optimal Solution	103
5.5.2	Heuristic Solution	104
5.6	Handover Design Protocol	106
5.7	Simulation Setup	109
5.8	Performance Metrics	111
5.8.1	Transmitting Power	111
5.8.2	MOS	111
5.8.3	Handover Dropping Probability	111
5.9	Simulation Results	112
5.9.1	Transmitting Power	112
5.9.2	MOS	113
5.9.3	Handover Dropping Probability	113
5.10	Discussion	114

6	Conclusions and Future Work	116
6.1	Conclusions	116
6.2	Future Work	118
A	Description of Parameters Used in Chapter 3	119
B	Description of Parameters Used in Chapter 4	120
C	Description of Parameters Used in Chapter 5	122
	References	124

List of Tables

2.1	Comparison of WiMAX and UMTS	22
3.1	Mobile WiMAX Applications and Quality of Service	41
3.2	WiMAX and UMTS Network Parameters	47
4.1	Delay Values for Codec	71
4.2	Packet Loss Values for Codec	72

List of Figures

1.1	Next Generation Integrated Network	3
2.1	UMTS System Architecture and Interface	16
2.2	UMTS Architecture	18
2.3	WiMAX System Architecture	19
2.4	WiMAX Protocol Stack	21
2.5	LTE Architecture	23
2.6	Tightly Coupled Architecture of UMTS and WiMAX	26
2.7	Loosely Coupled Architecture of UMTS and WiMAX	27
2.8	General Handover Process	29
3.1	Voice over Internet Protocol	36
3.2	Frame Structure of WiMAX	41
3.3	VoIP in WiMAX	44
3.4	VoIP in UMTS	45
3.5	Tendencies of MOS with regards to number of connections	52
3.6	Packet End-to-End Delay in UMTS	54
3.7	Packet End-to-End Delay WiMAX	55
3.8	Jitter in UMTS	56
3.9	Jitter in WiMAX	57
3.10	Packet Delay Variation in UMTS	58
3.11	Packet Delay Variation in WiMAX	59
4.1	RSS vs Distance	66
4.2	R-score vs Delay	71
4.3	Tightly Coupled Architecture for an Integrated Scenario	73
4.4	MOS vs Delay	75
4.5	Handover Design Protocol	81
4.6	MOS-Based Scheme	87
4.7	RSS-Based Scheme	88
4.8	Handover Dropping Probability	90
4.9	Call Dropping Probability	91
5.1	Energy saving at different layers	95
5.2	Loosly Coupled Architecture for an Integrated Scenario	99
5.3	Transmitting Power for Optimal vs Heuristic	107
5.4	MOS for Optimal vs Heuristic	108

5.5	Handover Design Protocol	109
5.6	Transmission Power	112
5.7	MOS	114
5.8	Handover Dropping Probability	115

Chapter 1

Introduction

The enormous success of the Internet has led to a great revolution in telecommunication. Today, telecommunication has become an indispensable part of everybody's day-to-day activities. For instance, services such as video on demand, music download, video streaming, video conferencing and VoIP are becoming part of users' daily activities. These applications demand high quality service, particularly for voice and real time sessions. Consequently, many wireless technologies, such as 3rd Generation (3G), Worldwide Interoperability for Microwave Access (WiMAX), Wireless Fidelity (WiFi), Universal Mobile Telecommunications System (UMTS) and Long Term Evolution (LTE), are emerging to satisfy the users' growing requirements to provide anytime and anywhere access to the Internet. In the recent past, the key focus in the telecommunication area has been on the Next Generation Wireless Networks (NGWNs). These networks will comprise different network access technologies to offer seamless access to the Internet, global roaming, high speed and higher user satisfaction.

The NGWN is a combination of the circuit-switching and packet-switching networks which will provide diverse services by maintaining the required Quality of Service (QoS). The NGWN is a wireless network which can cover a larger area, where a user can take advantage of different radio access networks to obtain better call quality. The service-related functions will be independent of the underlying transport-related technologies (Lee and Morita, 2006). It can be defined as a complex, easily available, convenient, converged, economical, effective, flexible, personalized, real-time, reliable and secure network (Korotky and Pfeiffer, 2009). Further, it will provide more applications than those in 2G, 3G, and 3rd Generation Partnership Project (3GPP), and will improve the quality of service. Unlike 3G, 2.5G and 2G which are largely based on the cellular network, the NGWN will support cellular, packet-based, satellite-based, and

wired networks with a backbone of core Internet Protocol (IP) network. The NGWN with multiple network technologies offers interactive gaming applications, multimedia services and many other real time applications to the users. In addition, it provides the luxury of utilizing the best available wireless technologies for different services to the users, enterprises and business organizations. In contrast, current homogeneous networks find it hard to provide a rich user experience with growing online users, social network and multimedia applications. However, the NGWN being the integration of different networks and each of its network technologies having different methods to provide QoS, the main challenge in the next generation network will be handover of a call from one radio access technology to another providing end-to-end QoS.

1.1 Preliminaries

QoS can be defined as a service which a network provides to reduce the packet loss. It provides assurance to the network for an end-to-end service and overall system performance. The objective of the QoS is to provide guaranteed quality to the users in terms of latency, throughput and available resources. It ensures a user receives a desired service for the applications. QoS categorizes the applications based on its importance and priorities. It plays a prominent role in real time applications such as watching online Television (TV). In these applications, the mobile network must provide reliable quality so that the users can watch the TV without any service disruption.

Network operators strive extremely hard to provide good and consistent QoS for the diverse user applications. Applications such as streaming, gaming and Internet Protocol Television (IPTV) require maximum resources and priority over other applications such as data download and email. In addition, another important issue is compatibility with the existing and previous generation networks considering these networks will take a long time to be replaced by the NGWN. Furthermore, mobility is one of the biggest problems in the NGWN. In order to provide mobility and better call quality experience to the terminal, the NGWN supports handover between different radio access networks for the moving terminals. The handover is a process of transferring the ongoing call from the serving base station to the new base station. This ensures the quality of the call does not degrade and also ensures the call continuity. Either in its idle or active mode the user will get attached to its home network base station. Moving terminals require a handover mechanism in the network to maintain the call continuity in the network when a user moves out of the coverage area of the current serving base station

or its home BS. The mobile terminal moving out of the home network will be attached to the new network called the visiting network. During the process of handover, there is the possibility of call rejection or delay in packet transmission for the ongoing session of the call, which might lead to poor QoS. Innovative technologies and extensive research in the area of mobility and QoS will make the NGWN work efficiently any-time, anywhere and provide seamless mobility to the users. Currently most of the mobile networks are circuit-switching. Changing the network from circuit-switching to packet-switching while supporting backward compatibility is not an easy task. It requires huge investment from the operators to provide the required infrastructure for the NGWN and also to upgrade the existing systems.

The NGWN migration will take place sooner, and the challenges such as compatibility with existing technologies, providing cost effective services to the users, mobility and many other issues are to be considered. As seen in Figure 1.1, the next generation network will be a convergence of cellular, packet, satellite and wired networks. The mobile terminal will have unrestricted access to various radio access technologies. The NGWN user terminals will have an option to select the best available technology depending upon the users requirements.

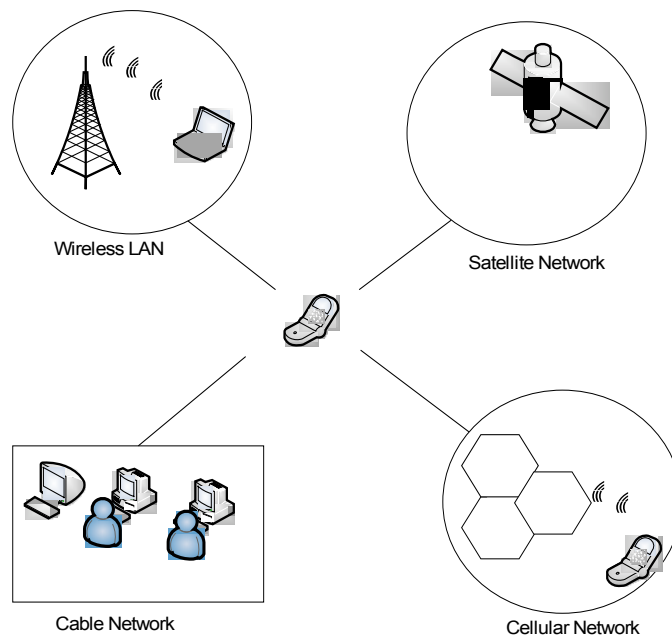


Figure 1.1: Next Generation Integrated Network

1.2 Challenges in the Next Generation Wireless Networks

The key technical challenges for the NGWN include, end-to-end QoS, mobility management, energy efficiency, call admission control, resource allocation, security and intelligent billing. In this section, we discuss some of the challenges involved in the NGWN. Comprehensive research in these areas will make the NGWN work efficiently. This thesis considers the QoS, mobility management and energy efficiency challenges in NGWN.

- QoS: The NGWN is a multi-network carrier with the mobile terminal roaming in different access, core, and back-bone technologies. For this reason it is very difficult to provide a user with guaranteed end-to-end QoS. Hence optimal QoS solutions are required to ensure guaranteed QoS in a sophisticated wireless domain. Further, the quality of applications is differentiated on an individual application basis, thereby adding more complexity to the system architecture. Applications such as video streaming and teleconferencing require more resources than the non-real time applications. Also, these applications are managed with different reserved resources at radio access, core network and the unmanaged Internet. Accepting these applications or calls without considering the QoS requirement might affect the perceived call quality. The existing system cannot guarantee the end-to-end quality for multimedia applications. This has encouraged many researchers to develop the QoS models in the NGWN (Li, Hamdi, Iang, Cao and Hou, 2000). In the case of a homogeneous network, the call has to be transferred from one cell to another, which has similar radio access technology. But in the case of NGWN, which is a heterogeneous network, the call can be transferred from one network to another which has a different radio access technology. For such a heterogeneous network with different radio access technologies, it is particularly challenging to provide the end-to-end QoS as a user will not be aware of this movement and the available radio access networks. A user terminal and carrier networks have to share the responsibility of providing excellent QoS to a user. Another critical challenge in NGWN is the radio access and core network of different technologies, such as UMTS, WiFi, WiMAX and LTE which have their own model of QoS stack in compliance with their respective standard bodies. Therefore, the NGWN will need a new mechanism to map the QoS of different access networks.

- **Mobility Management and Seamless Connectivity:** Mobility management and seamless connectivity allow a user to move among the multiple radio access technology networks. It comes in the scenario, where a user wants to switch the network for the reasons of quality and cost or in the scenario where a user is moving away from the current base station and has weak signal strength to support the ongoing session. Under such circumstances, the current base station cannot support the ongoing session, and it has to handover the session to the new base station (Akyildiz, Xie and Mohanty, 2004). The handover in the heterogeneous network is more complex and challenging, since the mobile terminal may be required to handover the call to a new network with a different radio access technology and a different core network. When a user moves to a different network, there is an immense challenge posed by the new network to admit or reject a user. Accepting a call without considering network parameters might affect other users. Mobility management plays a pivotal role in improving call quality, network congestion and access speed due to many access networks in the NGWN, Future NGWN mobile terminals will conveniently be able to attach to different access technologies depending on their application type. Due to many available networks, selection of the appropriate network in the NGWN is a challenging issue. There are several factors which can be considered, such as whether the application is video/voice/data, acceptable QoS for the application, cost of the radio access technology, access speed and more. If multiple radio access technologies are available, it is difficult to determine which radio access technology the call should be forwarded to and the criteria for selecting radio access technologies.
- **Energy Efficiency:** Since both energy cost and network electrical requirements show a continuous growth, energy efficiency is of great interest to the research community. In wireless communication, most of the network equipment operates on battery power which is a limited resource (Bolla, Bruschi, Davoli and Ranieri, 2009). There are many areas in the NGWN where more energy may be required than the traditional networks. Firstly, multimedia applications play a prominent role in energy consumption in the NGWN. These applications require much more energy than normal applications, such as voice. Secondly, the NGWN requires frequent switching between the different networks which consume enormous energy. Energy efficiency problems in wireless networks are addressed at various levels, such as network level, device level and application level. There is a trade-

off between the QoS and the energy efficiency. As a result, it is necessary to maintain a proper balance between them, therefore energy usage in the network should be managed efficiently as one of the essential and critical resources.

- **Call Admission Control (CAC):** The main objective of the CAC is to reduce the connection dropping probability for new and handover calls. The number of connection requests dropped by admission control to the overall connection requests in the network is the dropping probability. A call can be initiated in two scenarios, during handover and a new call. Depending upon the congestion in the network the call might be admitted or rejected or delayed. CAC in wireless communication is a process of administering the traffic volume. The decision of whether a new call has to be admitted, delayed, dropped or forwarded to a neighboring network is decided by CAC sub-system. The incoming traffic can be divided into real time (video and voice) and non-real time (images, text) (Chen-Feng, Liang-Teh and Der-Fu, 2011). The real time applications are extremely sensitive to delay, packet loss and jitter as compared to the non-real time applications. CAC plays an extremely influential role for the real-time applications.
- **Resource Allocation:** Resource allocation is a mechanism where an admitted call is allocated with the required resources, such as bandwidth and buffer. Allocating the required resources depends upon the application (voice, data, or video). Incoming traffic, whether it is data, voice or multimedia, can be admitted only if there are sufficient resources in the network. This requires a smart resource allocation scheme.
- **Intelligent Billing and Cost:** The NGWN can access or utilize services and resources beyond the service provider to which they belong. This requires an advanced customer management and billing system in place to support the new technologies that emerge in the NGWN. The billing system should be able to manage and generate a bill for users irrespective of the network they use. So efficiency of the system, to manage a vast customer database and generate bills, might pose a significant challenge for service providers. Service providers are working to extend the existing systems to provide flexible and advanced charging according to the pricing policy with customized billing (Hwang, Hwang, Ku and Chang, 2008), which is a challenging task. Cost is another issue in selecting radio technology. A user may be using a service with lower cost, but when he moves to another network the cost might not be similar. Selection of radio technology

based on available resources and cost is another concern in the NGWN.

- **Security and Fraud Management:** The open and distributed nature of the convergent NGWN architecture enables easy access to services, information and resources together with constant abuse by hackers, fraudsters and organized crime units (Bella, Olivier and Eloff, 2005). User identification based on the IP layer can be easily tampered with. The packets sent over the network can be easily marked with a "borrowed" IP address, enabling unauthorized users to impersonate legitimate ones. These intruders abuse services and benefit at the expense of the legitimate users, who are often unsuspecting until the bill arrives (Ericsson, 2004). The fraudsters can obtain a valid electronic serial number and mobile identification number during the registration process of the call. They can duplicate the same number on the other handset and utilize the services in the name of real user. The model (Bella, Olivier, and Eloff, 2005) discussed fraud management, detection and prevention techniques for the NGWN.
- **Location Registration:** In the current cellular network, the location update has to be maintained at two different locations namely Home Location Register (HLR) and Visiting Location Register (VLR). In (Lee, Lee and Cho, 2003) a mobile node uses both Mobile IP and Session Initiation Protocol (SIP) for providing mobility. The redundancy of having a separate registration for Mobile IP and SIP is an overhead. Integrating mobility management in mobile IP and SIP is an acceptable solution authors proposed. In the same way, there should be a centralized location updates for all the networks involved or centralized updates for each of the core networks in the NGWN.

1.3 Motivation

The major issue in the NGWN is guaranteeing the QoS and successful handover of the call across the networks whenever required. This demands the network reliability, low power consumption, bandwidth, timeliness, jitter, fault tolerance and seamless mobility among heterogeneous access networks. These needs are driving factors for the NGWN. VoIP in particular will be the most popular application in the future. VoIP demands high QoS for the moving user. In this regard, many problems exist to make the NGWN fully efficient and compatible with the existing services and technologies. For the moving user providing desired service

such as VoIP, by maintaining the acceptable QoS considering different networks and global roaming is a difficult task. The NGWNs unconventional needs demand efficient and reliable access technologies to satisfy users, as well as network providers. When a user initiates a call, there might be more than one radio access technologies available for a user, so the selection of an appropriate network for the ongoing call in the NGWN is an important issue. There are several factors which can be considered, such as:

- Does the application handle video or voice or data?
- What is the acceptable QoS for the application?
- What is the cost of the radio access technologies and access speed?
- Which is the best available radio access technology for call forwarding between the available networks?
- During handover process which base station to select from the available set of base stations in order to minimize the transmission power?

QoS is directly related to the quality of the voice or video that a user experiences. Users can be of various types depending on their personal needs, business needs and health-care needs. Lately, there is a trend of school graduates accessing tutorials online from their personal devices and listening to video tutorials available from the library resource. In the future, classroom lectures may go online, and students will be able to access them anywhere and interact with the lecturer. The health-care industry in particular is a thriving market. NGWN can provide different levels of health services to clients/patients such as remote health monitoring while they work at the office or stay indoors at home. Developed countries such as Japan are already providing these types of services to clients and looking to expand the network to offer services to more clients. In all these applications QoS plays a significant role as there is zero tolerance for data degradation in these services.

The seamless architecture requires the integration of technologies, such as Code Division Multiple Access (CDMA), General Packet Radio Service (GPRS), Global System for Mobile (GSM), Digital Subscriber Line (DSL) and more networks that can be connected to IP networks through different gateways (Motorola, 2005). These radio access technologies have different criteria for providing QoS. In this regard, it is challenging to provide a seamless service in the NGWN. The seam-

less service to the moving user is provided by using the handover technique. The main objective of handover is to continue the ongoing session of the call in order to provide an end-to-end QoS to a user. Seamless service provides the users with the same or better QoS when they move from one network to another or between different access points of the same network. Compared to previous decades the population of mobile users has increased beyond expectations. The current generation of the population is more familiar with multimedia applications. Considering these applications it is hard to provide an uninterrupted service to all users.

There are various mechanisms to support QoS for handover, such as reserving the bandwidth in the visiting network, or borrowing the bandwidth from the neighboring network/cell. Several mechanisms such as over-provisioning have been proposed in the past, but they are not particularly feasible as bandwidth is not utilized to the optimum level every time. Accommodating more bandwidth might solve the problem, but it is not a convenient solution as bandwidth is exceedingly scarce resource. Guaranteeing QoS is difficult due to the heterogeneous nature of networks and limited radio resources.

Handover comes in a scenario when a user is moving from the one network to another. For instance, if a user is moving out from the cellular network and he has a choice of attaching to WiFi or WiMAX. The choice depends on the availability of resources in these networks and which can provide a high quality of service taking cost of the service into consideration. To decide which network to select when a user is moving out of the current network in order to avoid jitter and reduce dropping probability is a challenging task. In this case, the network should determine whether a user is going to move from one network to another. It should transfer the required information to another network before a user goes into the new network and at the same time buffer the ongoing session packets. In case of any problem, handover process should restart from the point where it broke down. Diverse requirements with diverse applications on diverse networks are formidable challenge for NGWN. There are fuzzy logic and neural network mechanisms that might be used to solve this problem, but these are complex and incur delay of the handover process.

Another problem in NGWN is power management. Since the battery power is extremely limited for the mobile terminal it has to be used efficiently. It is therefore

necessary to have power saving mechanisms to extend battery life. There is a distinguishable amount of power spent during the handover process. To perform the handover the mobile terminal first needs to scan all the available channels which consumes a considerable amount of power. Moreover, handover may be dropped due to some unsatisfied constraints such as capacity constraint at the base station. To continue the handover process, channel scanning must be re-done thus consuming a significant amount energy at the mobile terminal. Another major portion of energy is consumed by wireless data communication. During the handover process, the mobile terminal may choose a base station which is required to transmit at a higher power level to guarantee satisfactory communication quality, thus consuming a significant amount of energy. The handover techniques play a prominent role in saving energy at the mobile terminal, so an energy-efficient handover solution should be implemented to minimize energy consumption for both channel scanning and data communication.

1.4 Contributions

In this thesis, the main focus is on providing end-to-end QoS to a user during the handover process for the NGWN. Since UMTS and WiMAX are among the main wireless networks in NGWN, QoS in UMTS and WiMAX networks will be discussed first. Then quality-based handover scheme for the integrated UMTS-WiMAX network is proposed. This handover scheme is further enhanced to provide an energy efficient handover maintaining acceptable quality for a user.

In order to achieve the above goals, in the first part of this thesis, extensive simulation has been conducted to evaluate the performance of WiMAX and UMTS for supporting VoIP traffic using OPNET. Application classification and QoS restructuring for different networks are done and the behavior of different applications on different networks is presented. In the future when UMTS and WiMAX allow users to select any of the available networks, these classifications of QoS will make it easier for the network operators to allow a user to switch to the network that is best suited for the real-time application. We believe that the classification of different QoS requirements from the real-time multimedia applications will help to select the best available network without degrading the QoS of the applications. Several important critical parameters, such as MOS,

end-to-end delay, jitter and packet delay variation are analyzed. This study is the first step towards exploring possible implementations of the NGWN.

In the second part, a novel handover scheme compliant with the IEEE 802.21 standard that enables a wireless access network to transfer the call between cells or networks has been proposed. This scheme takes care of the quality of the call and load among all the available attachment points. The base station selection problem has been formulated as an optimization problem with the objective to maximize the call quality. A scheme is presented to forward data packets to the most appropriate attachment point in order to maintain good call quality. In order to find effective solutions, extensive simulation work has been conducted using a scenario of urban network environment with VoIP call in a WiMAX and UMTS integrated network. Critical QoS parameters like MOS, CDP and HDP are analyzed. The integrated network of UMTS and WiMAX is simulated using MATLAB. A scenario is created where base stations of the UMTS network and the access point of WiMAX networks are deployed. These users move randomly, and when the QoS of a call goes down they initiate a handover process. The proposed scheme is compared with the Received Signal Strength (RSS)-based handover scheme. Results show that the proposed scheme provides higher MOS values thus improving the perceived quality of the call and improving the Handover Dropping Probability (HDP) and Call Dropping Probability (CDP). It is a QoS aware scheme which guarantees the call quality to a user.

In the third part, the problem of minimizing energy consumption at the mobile terminal side through optimal handover with a guarantee on the communication quality is investigated. The issue of designing an energy-efficient handover scheme to minimize energy consumption at the mobile terminal subject to the constraint on communication quality is addressed. The energy consumption for both data communication and channel scanning is taken into account. To save energy for wireless data communication, the minimum transmission power needed for each channel to provide the desired QoS is computed. For channel scanning, the handover dropping probability is used to estimate the energy consumed by scanning. By formulating the handover as an optimization problem, both a centralized solution and a heuristic solution for base station selection is proposed. An energy-efficient handover protocol is then designed based on the IEEE 802.21 standard. Simulation results show a substantial improvement in terms of call dropping probability, power consumption and MOS when compared to the tra-

ditional RSS-based handover technique.

1.5 Thesis Outline

The rest of this thesis is organized as follows:

Chapter 2 begins by introducing different radio access network technologies. Then several widely used networks, which are contenders for NGWN, are discussed. The architectures of UMTS and WiMAX are discussed in detail as these two networks are used for my further study. The comprehensive knowledge of these networks which is required for understanding my research is presented. An overview of the tightly coupled and loosely coupled architectures in NGWN is discussed. Lastly, an overview of handover process is given.

Chapter 3 reviews related work in the areas of QoS in UMTS and WiMAX. The QoS models of these networks are discussed. The simulation set-up is then presented, where we discuss the UMTS and WiMAX simulation model using OPNET and the simulation configuration. In this study, both models are simulated in different scenarios. The metrics for performance evaluation used in this study are discussed. In this chapter, we have carried out extensive simulations to evaluate the performance of UMTS and WiMAX.

Chapter 4 addresses the handover problem in NGWN. Firstly, the novel quality-based handover scheme which is compliant with the IEEE 802.21 standard has been proposed. The Mean Opinion Score (MOS) has been used which is the function of delay and packet loss as quality parameter. The optimal base station selection problem has been formulated to maximize the quality of the VoIP call. An integrated scenario with UMTS and WiMAX networks is considered. The solution for selecting the optimal base station is presented. An analytical model is derived for evaluating the proposed scheme. Secondly, the design of handover protocol is described, where the parameter acquisition and the detailed design principles of handover protocol are discussed. Finally, a simulation model for MOS-based handover algorithm in MATLAB is discussed. The proposed scheme is then compared with the RSS-based scheme. The results of both schemes are presented to evaluate the performance.

Chapter 5 focuses on the handover problem in NGWN by taking energy efficiency criteria into consideration. The system model for the handover problem is

presented. In this chapter, a novel handover scheme which provides optimal quality and minimizes energy consumption of the mobile terminal is proposed. The problem is divided into two parts: energy in communication and energy during scanning. The solution for the selection of the optimal base station is presented by proposing two schemes - heuristic and optimal. An analytical model is derived for evaluating the proposed scheme. The results of both schemes are compared. Finally, the simulation model is discussed, and the results are presented.

Chapter 6 concludes the thesis and provides a brief discussion of future work.

1.6 List of Publications

- Performance Evaluation of Quality of VoIP in WiMAX and UMTS, Sheetal Jadhav, Haibo Zhang, Zhiyi Huang, in Proceeding of the 12th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2011, pp.375-380, ISBN 978-0-7695-4564-6, Best Paper Award.
- Vertical Handover Algorithm for next generation networks, Sheetal Jadhav, Haibo Zhang, Zhiyi Huang, in 26th International Conference on Advanced Information Networking and Applications (AINA), March 2012, doi: 10.1109/AINA.2012.105, pp.479-486.
- Quality of VoIP in WiMAX and UMTS, Sheetal Jadhav, Haibo Zhang, Zhiyi Haung, in Journal of Internet Technology, May 2012, Vol. 13, No. 3, pp.491-500.

Chapter 2

Background

Wireless networks have gone through revolutionary changes in the last few years due to their increasing demand. With the rapid growth of wireless packet-switched networks, sending data through the Internet rather than the Public Switched Telephone Network (PSTN) has become a better option in terms of reducing cost for both users and service providers. Mobile phone users can make voice/video calls through the Internet with better communication quality and less cost than using PSTN.

An attractive wireless technology for VoIP is the WiMAX specified by IEEE 802.16 standard, aiming to provide wireless access over long distances in a variety of ways from point-to-point communication to mobile cellular access. WiMAX provides a wide coverage area with lower cost of network deployment. The coverage area of a single WiMAX cell is around 30 to 50 km, and its speed is up to 40 Mbps (Chakraborty and Bhattacharyya, 2010). Moreover, WiMAX supports QoS by providing different service classes for both real-time and non-real-time traffic. Thus, WiMAX is an immensely attractive technology for providing integrated voice and video services for VoIP.

Another emerging wireless technology is the development of UMTS as a part of 3G network. UMTS has circuit-switched transmission for voice and packet-switched transmission for text, video, digitized voice and multimedia. As a complete network system it supports high mobility to fulfill a user demands in any places including office, home, urban and rural areas. UMTS supports packet-switched applications including real-time multimedia applications, such as VoIP with a peak down-link data rate of 14.4 Mbps (QUALCOMM, 2008).

The rest of the chapter is organized as follows. Section 2.1 briefly describes the architectures of UMTS, WiMAX and LTE. Section 2.2 describes the integrated architecture and section 2.3 gives an overview of the handover process.

2.1 NGWN Access Technologies

In future, homes and businesses will require high speed Internet connections with data rates in the range of Gbps. To meet this requirement new radio access technologies with better performance are being launched. In this section, we will discuss several leading radio access technologies in the market.

2.1.1 UMTS

2G systems were designed originally for voice communication. GSM is an immensely popular second generation cellular telecommunications system which is commonly used worldwide. Recent developments in mobile communication have changed the way people communicate. New technology like UMTS, which is 3G is becoming more popular, due to the speed and support of multimedia applications it provides along with the voice communication.

The 3G system has evolved from GSM, which was initially developed for voice communication. 3G provides wireless service for both data and voice communications. Its popularity has increased due to the support for multimedia applications. The 3G system utilizes much of the GSM infrastructure for the voice call. It uses GPRS for Internet-based service, which has data connection with higher bandwidth and is a packet-switched wireless system (ETSI, 1998). GPRS utilizes the packet-switched infrastructure and provides packet data service by adding two new types of nodes to the network: Gateway GPRS Support Node (GGSN) and Serving GPRS Support Node (SGSN). Through GPRS, GSM is evolving into the 3G cellular network called UMTS. UMTS, while reusing the GSM/GPRS core network, has an entirely different radio access network employing Wideband Code Division Multiple Access (WCDMA), instead of the Time-division multiplexing (TDMA) used in GSM/GPRS (Agharebparast and Leung, 2002).

In 1999, the 3rd Generation Partnership Project (3GPP), which is a collaboration between groups of telecommunication associations, launched UMTS, the first

3G release as part of the International Mobile Telecommunications-2000 (IMT-2000) family of 3G standards (Holma and Toskala, 2000). The 3GPP is a forum (3GPP, 2006) where standardization is handled for High Speed Downlink Packet Access (HSDPA) and High-Speed Uplink Packet Access (HSUPA). It has also been handled for the first Wide-band Code Division Multiple Access (WCDMA) specification release. HSDPA was standardized as part of 3GPP release. The HSDPA peak data rate available at the terminal was initially 1.8 Mbps. It eventually increased to 3.6 Mbps and 7.2 Mbps during 2006 and 2007 and potentially beyond 10 Mbps. The HSUPA peak data rate in the initial phase is expected to be 1-2Mbps with the second phase pushing the data rate to 3-4Mbps (Holma and Toskala, 2006a).

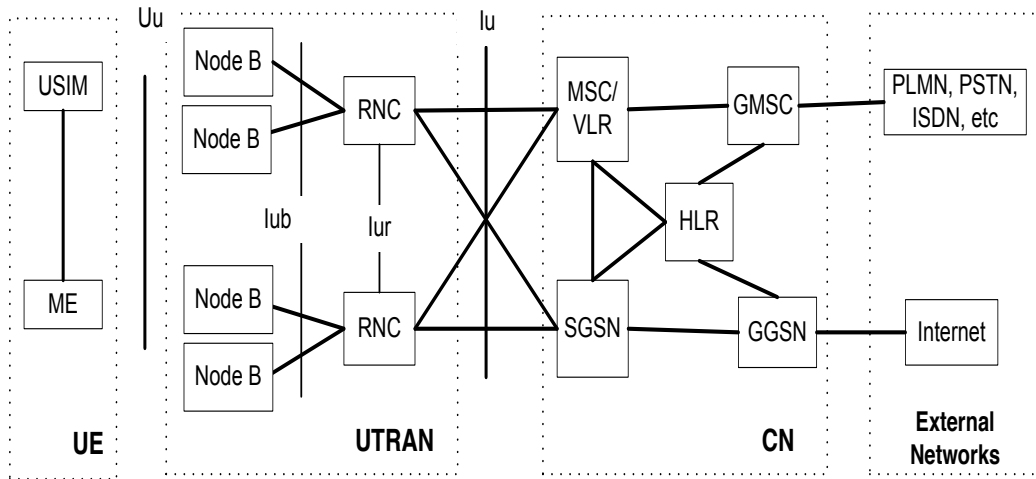


Figure 2.1: UMTS System Architecture and Interface

This section provides an overview of the UMTS network. UMTS supports two modes for channel access: Time division duplexing (TDD) and Frequency division duplexing (FDD) (Holma, Toskala, 2000). The UMTS architecture is shown in Figure 2.1. The UMTS architecture comprises of Core Network (CN), UMTS Terrestrial Radio Access Network (UTRAN) and the User Equipment (UE) (Lopes, 2008) (3GPP, 2005).

- **Core Network (CN):** The core network of UMTS has a Circuit Switched (CS) and Packet Switched (PS) domain. The circuit-switched domain is adapted from the GSM network and handles the GSM calls. The packet-switched domain handles IP-based calls. The main components of the CN are:
 - Home Location Register (HLR): HLR is a massive database that stores

information, such as a user identity, services to which a user is subscribed, billing information, user location and user preferences.

- Mobile Switching Centre/Visitor Location Register (MSC/VLR): VLR is a database for a user moving from one cell to another. When the handover of a call is done, a user information is temporarily transferred to VLR. It has a function similar to the HLR, and it is used for the moving user. MSC is responsible for switching voice and data connections in the CS domain.
 - Gateway MSC (GMSC): GMSC is a switch for connection to the external network on the CS domain.
 - Serving GPRS Support Node (SGSN): SGSN does not deal with any radio related functions. It is a switch for connection on the PS domain. It has a function similar to MSC. RNCs are connected to SGSN. SGSN performs the function of tunneling, mobility management and context activation. Single or multiple SGSN's are connected to the GGSN.
 - Gateway GPRS Support Node (GGSN): GGSN is a switch which is present on the packet-switched domain to establish connectivity with the external networks. The function of the GGSN is IP address assignment, authentication, IP address mapping and Packet Data Protocol (PDP) context. The packets coming from the SGSN are converted into the appropriate PDP format by GGSN and then sent to the respective packet network.
- **UMTS Terrestrial Radio Access Network (UTRAN):** Figure 2.2 shows the major building blocks of the UTRAN. The UTRAN is comprised of Node-B and Radio Network Controller (RNC). Each Node-B is connected to one RNC. RNC can have one or many Node-Bs connected to it. Node-B is responsible for communication with a user equipment which is through the radio interface. The connection is then transferred to CN through the RNC. UTRAN uses the WCDMA for the air interface.
 - Node-B: When a call is initiated the Node-B functions convert the data flows from the Uu radio interface to the Iub interface. It also performs modulation/demodulation, transmission/reception and Radio Resource Management (RRM).
 - Radio Network Controller (RNC): The RNC has three logic components: a) Controlling RNC (CRNC), which controls the logical resources of UTRAN

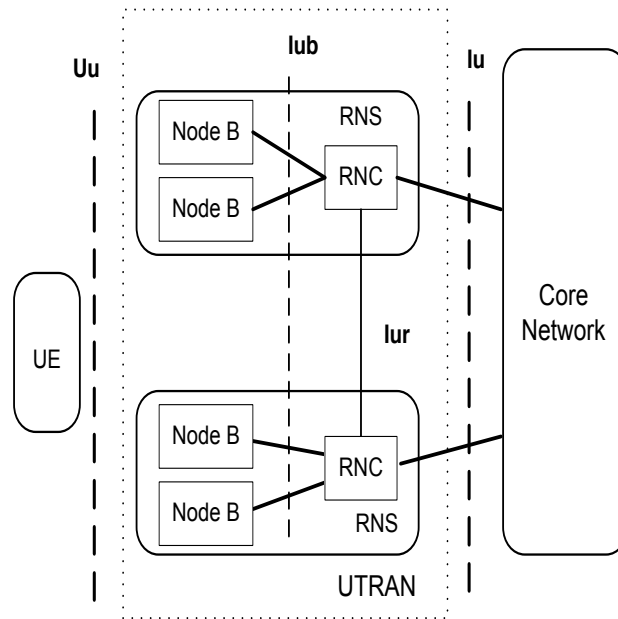


Figure 2.2: UMTS Architecture

access points; b) Serving RNC (SRNC), which ends the Iub and Iu interfaces, and c) Drift RNC (DRNC), which is any RNC other than SRNC controlling cells used by the mobile terminal. It performs functions such as macro diversity and splitting.

The Iur interface allows handover between NodeBs belonging to 8 different RNCs. The RNC manages the Radio Resource Control (RRC) and call admission control, where it decides whether to accept the call depending upon its capacity and support for the application. It manages the handover which UE initiates and takes the decision to handover. In collaboration with NodeBs, RNC performs Radio Resource Management (RRM), such as code allocation, channel allocation, broadcasting the signals, power control and packet scheduling.

- **User Equipment (UE):** The UE is composed of:
 - User Device (UD): The user device can be a mobile phone, laptop or a desktop, which can support GSM. It supports the voice call and/or GPRS, hence supporting the Internet based applications.
 - UMTS Subscriber Identity Module (USIM): It is a smart card that stores the subscriber identity, authentication information and encryption keys.

2.1.2 WiMAX

The WiMAX is an evolving IEEE standard and is also known as IEEE 802.16. WiMAX, like 2G/3G networks, can provide service on the scale of Metropolitan Area Network (MAN) with high bandwidth. The WiMAX wireless technology is called the last-mile solution for wireless broadband access. It can also act like a hot-spot. WiMAX has benefits in terms of spectral efficiency, wider coverage, easy deployment and frequency re-use. IEEE standard just provides the WiMAX technology. A large organization called WiMAX forum made of network operators, academics and telecommunication members work on the compatibility, technicality, regulatory, and marketing aspects of the WiMAX.

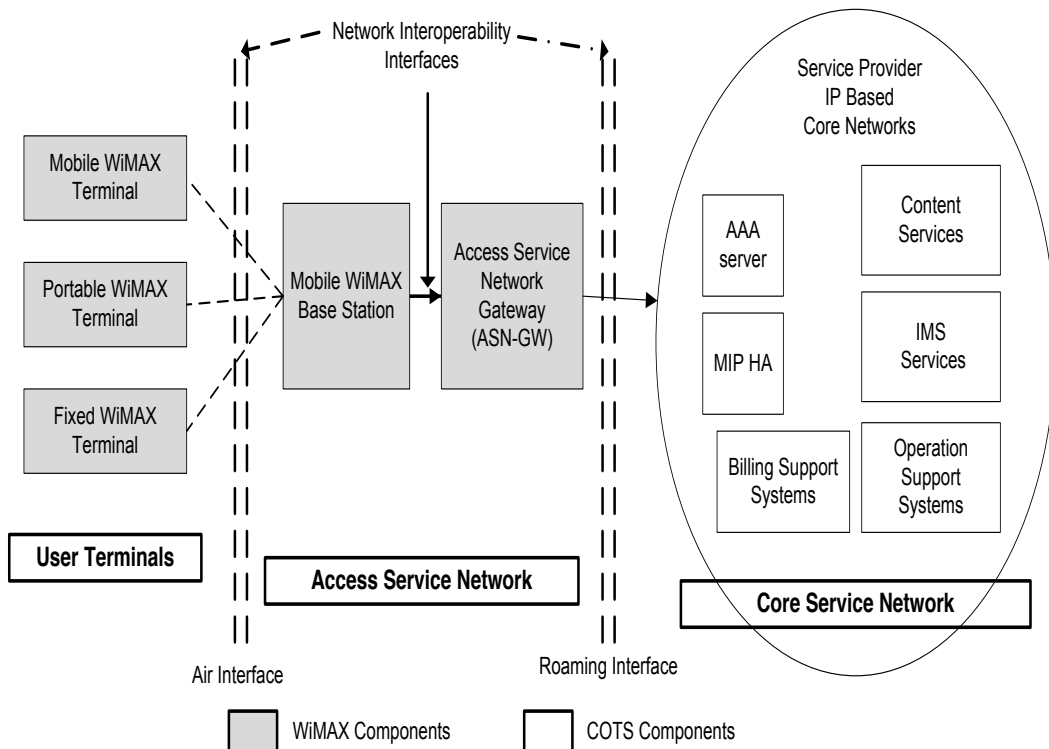


Figure 2.3: WiMAX System Architecture

Figure 2.3 shows the WiMAX architecture. The WiMAX architecture is comprised of the following components (Xu, Zhang and Zhou, 2007b):

- **User Equipment:** UE can be a mobile device, laptop or PC which uses the network to access the service.
- **Access Service Network Gateway (ASN-GW):** The ASN-GW typically acts

as a layer 2 for the WiMAX network. It consists of one or more base stations and one or more ASN gateways. ASN performs the function of radio resource management, admission control, management of the mobility tunnel with base stations, paging and Authentication, Authorization and Accounting (AAA) client functionality.

- **Connectivity Service Network (CSN):** Connectivity to the Internet and other public and private networks is provided by the CSN. It consists of application server, strategy agent and AAA server including application agents. The QoS strategy agent is located in the CSN which also performs the function of IP address management, support for roaming between different Network Switch Providers (NSPs), location management between ASNs, mobility and roaming between ASNs.

The IEEE 802.16 uses the Open System Interconnection (OSI) model. It defines and provides technical specifications for the lower two layers, physical layer and Media Access Control (MAC) layer. The MAC layer is a part of the data link layer. Figure 2.4 shows the layers of WiMAX. The difference in WiMAX PHY layer with respect to other technologies, such as UMTS is that WiMAX includes Orthogonal Frequency Division Multiplex (OFDM). In OFDM, the available bandwidth is divided into multiple frequency sub-carriers, which results into higher spectral efficiency. The other advantage of OFDM is that different QoS can be assigned to each of the users. WiMAX works well in Non Line Of Sight (NLOS) conditions, where the electromagnetic waves might not have a direct path between transmitter and receiver, but it takes several paths to reach the receiver. This is one of the reasons why WiMAX deployment is immensely popular in urban areas. This is done by utilizing advanced antenna diversity schemes and hybrid Automatic Repeat Request (hARQ) (Li, Qin, Low and Gwee, 2007a). WiMAX in its early stages just supported the line-of-sight (LOS) transmission.

One of the features of the MAC layer of WiMAX is that it is designed to differentiate services among traffic categories with different multimedia requirements (Cicconetti, Eklund, Lenzini and Mingozi, 2006). WiMAX offers some flexible features that can potentially be exploited for delivering real-time services. Though the MAC layer of WiMAX has been standardized, there are certain features that can be tuned for specific applications and channels. For example, the MAC layer does not restrict itself to fixed-sized frames, but allows variable-sized

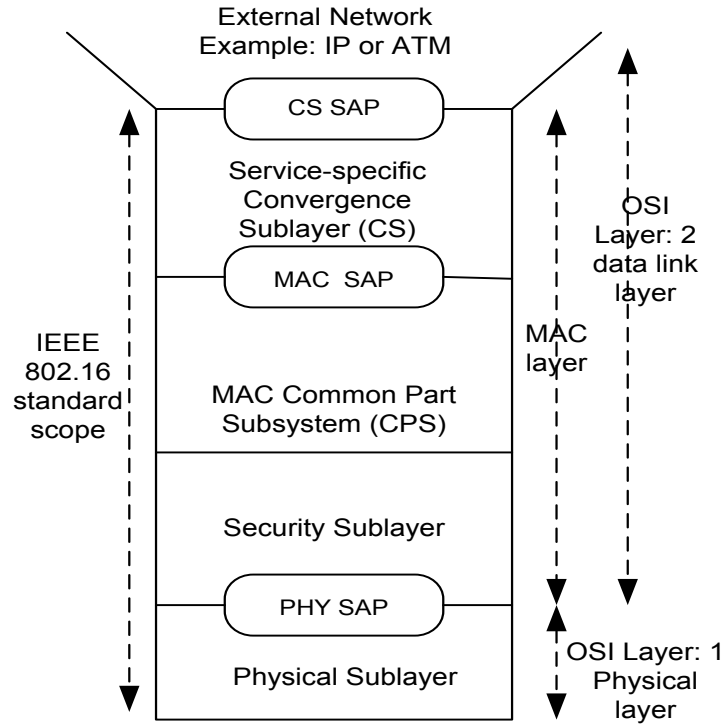


Figure 2.4: WiMAX Protocol Stack

frames to be constructed and transmitted. This is very useful for framing VoIP packets (Black, 1999; Ghosh, Wolter, Andrews and Chen, 2005).

When a call is initiated in WiMAX, depending upon the call type, the base station assigns a QoS class to the connection. This is not the case with UMTS which operates on the best effort QoS type. WiMAX supports transport technologies, such as IPv4, IPv6, and ATM to maintain compatibility with operators 17 transport technologies (Abichar, Peng and Chang, 2006).

Figure 2.4 shows the WiMAX MAC layer structure. The MAC layer consists of three sub-layers :

- Convergence Sub-layer (CS): which is service-specific and maps data from the upper layer to MAC Service Data Units (SDUs), which is then used by the MAC common part sub-layer.
- MAC Common Part Sub-layer (MAC CPS): performs the function of resource allocation, channel establishment and system access.
- Security sub-layer: which provides the authentication and authorization functionalities.

Table 2.1 shows the comparison of the two networks (Andrews, Ghosh and Muhamed, 2007).

Table 2.1: Comparison of WiMAX and UMTS

Parameter	WIMAX	UMTS
Peak down-link data rate	46Mbps with 3:1 DL-to-UL ratio TDD; 32Mbps with 1:1	14.4Mbps using all 15 codes; 7.2Mbps with 10 codes
Peak up-link data rate	7Mbps in 10MHz using 3:1 DL-to-UL ratio; 4Mbps using 1:1	1.4Mbps initially; 5.8Mbps later
Bandwidth	3.5MHz, 5MHz, 7MHz	5MHz
Modulation	QPSK, 16 QAM, 64 QAM	QPSK, 16 QAM
Multiplexing	TDM/OFDMA	TDM/CDMA
Duplexing	TDD initially/FDD	FDD
Frequency	2.3GHz, 2.5GHz, and 3.5GHz initially	00/900/1,800/1,900/2,100MHz

2.1.3 LTE

3GPP has undertaken the project of LTE standard following the success of High Speed Packet Access (HSPA). The objective of LTE is to support wireless communication with high-speed and spectral efficiency. The architecture of LTE is based on GSM and UMTS network technologies. LTE is also called EUTRA (Evolved UMTS Terrestrial Radio Access) as it is an extension of the UMTS architecture with enhanced features. In the LTE architecture, the main change is in the radio-access network, where the Node-B and RNC is replaced by eNodeB alone which can perform more functions.

LTE is designed to support advanced features. It uses the OFDM modulation technique to achieve a high data rate. The spectral efficiency is improved as it is based on a shared channel and does not have any dedicated channels to carry the data. It provides seamless mobility within the heterogeneous network

and supports end-to-end QoS. It also has inter-working architecture with existing UTRAN and provides a better handover quality.

Figure 2.5 shows the architecture for LTE.

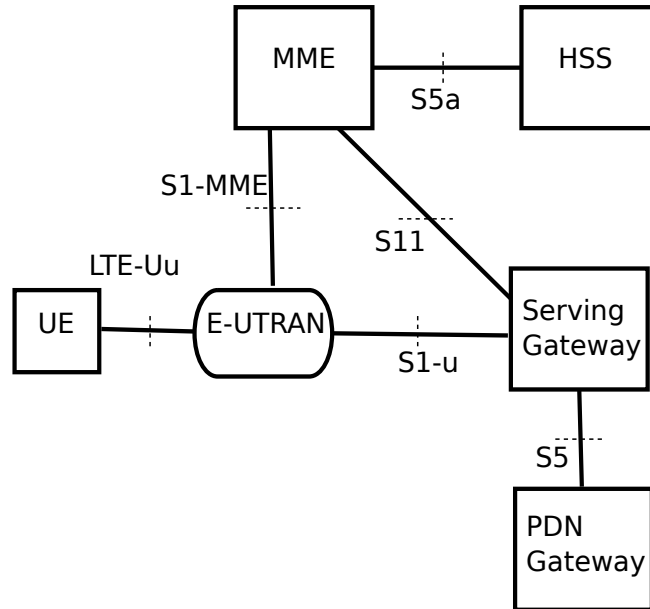


Figure 2.5: LTE Architecture

The architecture of LTE is broadly divided into following components:

- **Access Network:** Access Network of LTE is also called as E-UTRAN. It consists of eNodeB which performs the radio related functions. The eNodeB is connected to a user through the air interface, and performs function such as radio resource management, admission control and mobility management.
- **Core Network:** The core network of LTE is much more complex than the access network. The core network consists of:
 - **PDN Gateway (P-GW):** Responsible for IP-address allocation and QoS management. QoS is based on the traffic type.
 - **Serving Gateway (S-GW):** S-GW performs the administrative functions such as charging a user from the foreign area network. The IP-address assigned by P-GW to mobile terminal is transferred through the S-GW.

Apart from these two important modules, the core network consists of Home Subscriber Server (HSS) and Mobility Management Entity (MME).

2.2 Integrated Architectures

4G and beyond networks will demand integration of different wireless technologies. 3G wireless networks provide wide area coverage and dedicated support for mobility, whereas WiMAX provides a higher data rate, but has limited mobility. 3G and WiMAX have different QoS models; they have different characteristics in terms of protocols. Integration of 3G and WiMAX will allow a user to experience seamless connectivity, a faster data rate, selection of diverse radio access technologies, a reduced connection dropping probability for handover, better call blocking probability and improved utilization of bandwidth. Each user will access the service, which is subscribed with multiple access technologies, such as GPRS, UMTS, GSM, wireless local area network (WLAN), WiMAX and LTE (Nguyen-Vuong, Agoulmine and Ghamri-Doudane, 2007).

The integration of multiple radio access technologies plays an important role in the success of NGWN. In order to provide end-to-end QoS in the integrated networks, there should be a resource management and traffic controller in both the access technologies. Due to the diverse characteristics of QoS requirements, the system performance varies considerably, making QoS provisioning more difficult. To provide QoS continuity, it is critical to evaluate the availability of resources to satisfy the active session among the available radio networks. It is difficult to obtain the resource information and QoS constraints of neighboring networks from the perspective of the terminal equipment.

In this section, the existing integration schemes between UMTS and WiMAX are reviewed. The UMTS network is one of the most widely used networks, and WiMAX is one of the popular networks which provides a faster data rate and supports application like VoIP remarkably well. As a result, a user can choose the best available technology that suits his/her needs. A user can choose UMTS/GSM if he wants wider coverage, WiMAX for high speed and WiFi for lower cost. Vertical handover maintains the service connectivity for a user, when they roam between different radio access technologies. In this work, we use an integrated architecture between WiMAX and UMTS, considering their availability and wide deployment around the world. This work on the handover can be extended to any network, such as LTE and WiFi. The authors (Alamri and Akkari, 2012) reviewed the various inter-working architectures and handover scenarios between UMTS and WiMAX. The authors in (Sun, Stevens-Navarro and Wong,

2008) (Psimogiannos, Sgora and Vergados, 2010) proposed the vertical handover algorithm by taking connection duration, QoS parameters, mobility and location information, network access cost and the signaling load incurred on the network. The inter-working architectures are mainly classified into tightly coupled and loosely coupled architectures depending upon the coupling point.

2.2.1 Tightly Coupled Architecture

In the tightly coupled architecture, WiMAX and UMTS are integrated on the access level. In the tight coupling scenario, the WiMAX network may emulate a RNC or a SGSN (Bin, Martins, Bertin and Samhat, 2010). The Inter-Working (IW) sub-layer is introduced on top of the PDCP (Packet Data Convergence Protocol) sublayer of UMTS and MAC. In the article (Liu, Martins, Samhat and Bertin, 2008), the authors proposed the tight coupling architecture to achieve the interconnection between UMTS and WiMAX systems. Whenever handover of a call is made to a different radio access network, the IW sub-layer performs a function similar to the link layer control. In the case of tightly coupled architecture, since the networks are tightly coupled, the handover latency is reduced. In the tightly coupled architecture, the networks share the AAA and billing system, so the authentication of the call is completed more rapidly. These networks are relatively more secure than loosely coupled architecture. However, the complexity of the system increases and such an architecture requires changes in the core networks to handle the QoS and traffic routing. By definition in the tight coupled system, the subsystems of different radio access networks are linked together, and share the workload (Khan, Ismail and Dimiyati, 2010).

Figure 2.6 shows the tightly coupled integrated architecture between UMTS and WiMAX. From the figure, it is evident that RNC acts as a controller for both UMTS and WiMAX. When a user is in WiMAX network, the call is handled by WiMAX in the normal way. It is connected to the Internet through ASN-GW. The case is the same when a user is in the UMTS network. When a user wants to transfer a call from UMTS to WiMAX or vice versa, it is tunneled through the RNC which acts as an inter-mediator to transfer the call. Also, both the networks can have the same HA, AAA and billing system.

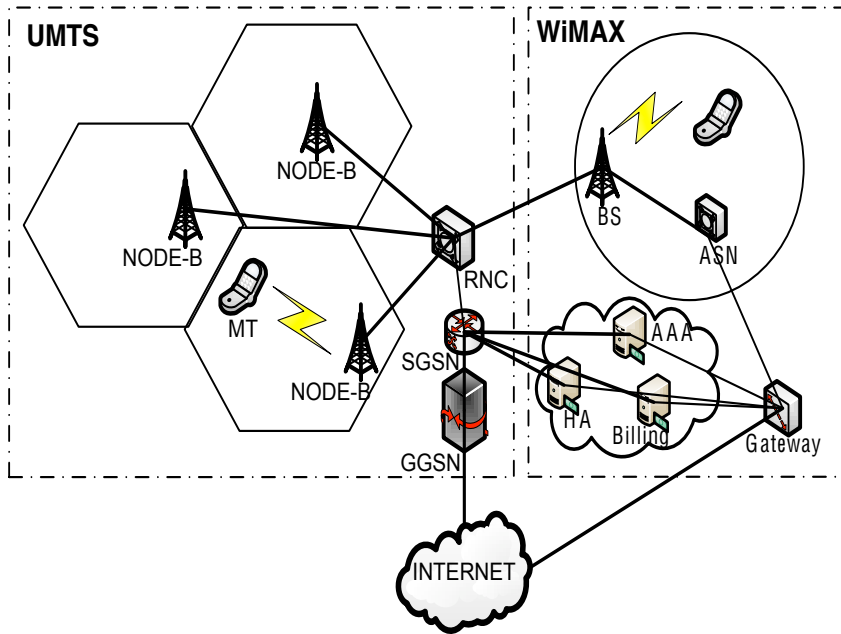


Figure 2.6: Tightly Coupled Architecture of UMTS and WiMAX

2.2.2 Loosely Coupled Architecture

Loosely Coupled Architecture is a system where subsystems of the different radio access networks work independently. The two radio access networks are inter-linked by a subsystem which acts as inter-mediator to forward the call and they do not share the workload. Each network has their own AAA and billing mechanism and operates through the intermediate handover controller. They are easy to deploy as they do not need many changes in the core network as networks operate independently of each other. In (Feder, Isukapalli and Mizikovsky, 2009), the authors proposed the loosely coupled architecture, where the networks have separate and independent data paths to the core network.

Figure 2.7 shows the loosely coupled integrated architecture between UMTS and WiMAX. From the figure, it is evident that both UMTS and WiMAX operate separately. The RNC of UMTS and ASN-GW of WiMAX is connected to the handover controller. The handover controller which manages the handover between these two networks. Whenever a user wants to transfer from UMTS to WiMAX or WiMAX to UMTS the session is tunneled through the handover controller which acts as an inter-mediator.

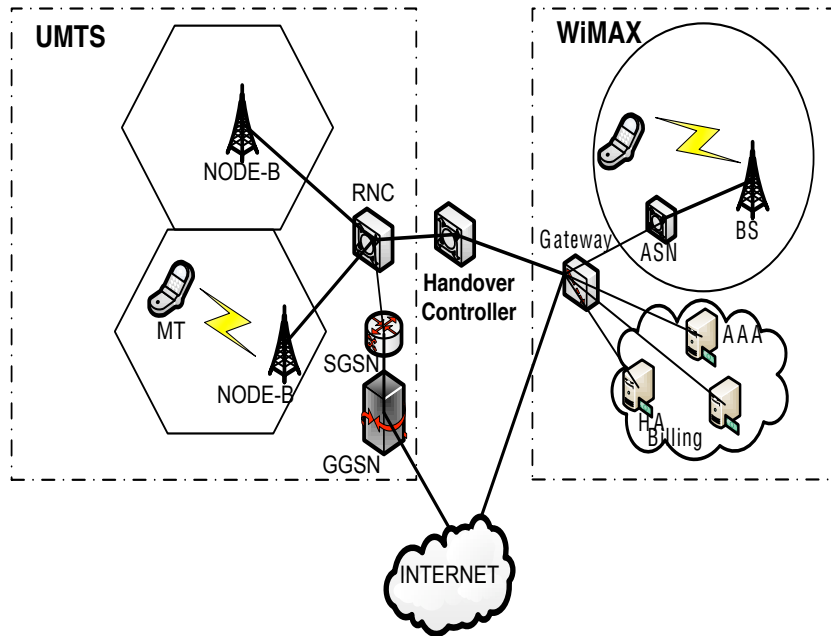


Figure 2.7: Loosely Coupled Architecture of UMTS and WiMAX

2.3 Handover in NGWN

2.3.1 Handover Overview

Seamless mobility is an important feature of the NGWN. In wireless communication, when a call is established, there is a dedicated radio link channel allocated to a user. Each user is associated with a single base station called the serving base station through which it is connected to the destination. The serving base station has limited coverage area and cannot handle the call if a user moves out of its coverage area. Therefore, it is required to transfer the call to another base station which can support the call. Handover is the process of changing the channel associated with the current connection while a call is in progress (Zeng and Agrawal, 2002). Handover maintains the call connectivity within the network by switching the call to another base station. It guarantees the connection to the moving mobile terminal.

The decision of making handover can be at the mobile terminal or at the network. To make a handover there has to be at least one base station available for the mobile terminal to hand the call over. The connectivity of mobile terminal is continuously monitored by the base station.

2.3.2 Handover Classification

In order to improve the efficiency of the network, a proper handover scheme has to be used when a user moves between different cells of the same or a different network. There are basically two types of handover depending on how the call is handed over from the current cell to a new cell of the same or a different network.

- Hard Handover: This type of handover is also called a break before make connection. In this type of handover, the mobile terminal communicates with only one base station at a time. When a handover takes place the current connection is broken, and a new connection is made. When the current connection is broken the resources associated with the old base stations are released, and the mobile terminal is allocated new resources associated with the new base station. A hard handover mechanism is not suitable for real-time applications as it might incur delay in switching the base stations. It is particularly suited for the delay-tolerant applications which are non-real time.
- Soft Handover: This type of handover is also called a make-before-break connection. Soft handover is a process of making a new connection before breaking the previous one. Thus, a user has two connections with the two different base stations. The soft handover is used for time critical applications which are not delay-tolerant. The UMTS and WiMAX networks support soft handover.

Traditionally the handover process was only within the cells of the same network. However, due to the growing number of networks and wireless access technologies, there was a need to transfer the call between cells of different networks. Handover is classified as horizontal and vertical handovers, depending upon whether the connection is transferred within the network or to a different network.

- Horizontal handover: When a connection is handed over to the base station within the same network, it is called horizontal handover or intra-system handover. Since both involved base station share the radio resources, network interface and QoS parameters, it is relatively easy to make a horizontal handover.
- Vertical handover: When the connection is handed over to the base station of a different network, it is called a vertical handover or inter-system han-

dover. The connection is transferred to a different radio access network, e.g. handover between UMTS and WiMAX. During vertical handover the IP-address of the connection is changed. It requires mapping of QoS parameter as both networks have a different QoS model.

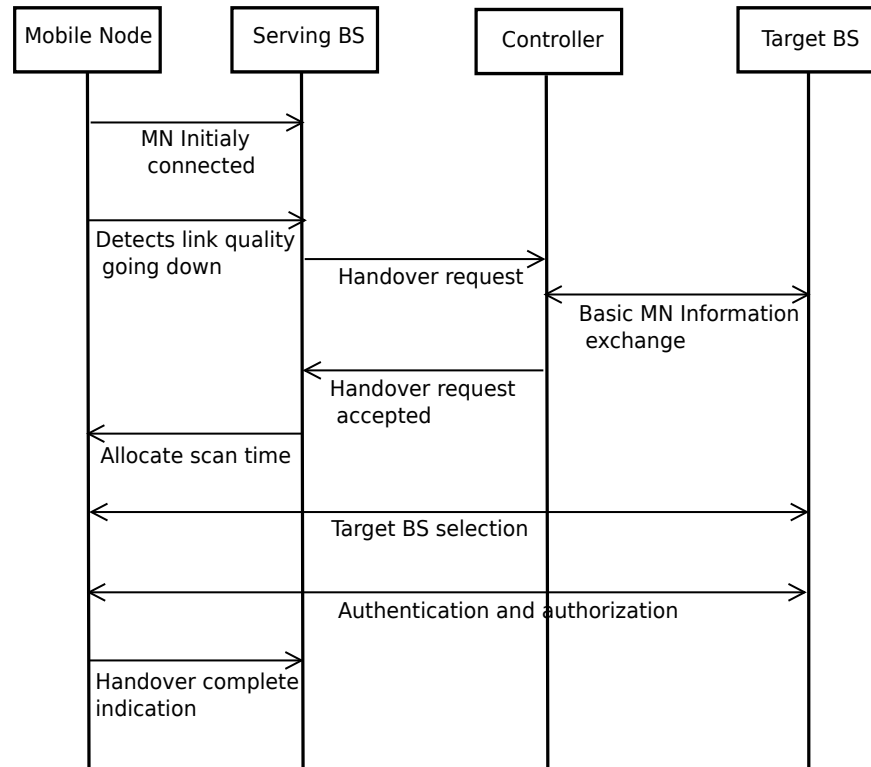


Figure 2.8: General Handover Process

2.3.3 Handover Process

The handover process starts from the time the handover is initiated until the decision is made.

- Handover Initiation: The handover process can be initiated by the network or the mobile terminal. In case of the network initiated handover, the mobile terminal periodically sends the report of the connection status to the controller and based on the report the controller decides if the handover needs to be performed. In case of the mobile terminal initiated handover, the mobile terminal takes the decision of the handover. The criteria for handover depends on the handover scheme which is used in the network. It can be based on QoS, Received Signal Strength (RSS) or SNR.

- **Network Discovery:** When a mobile terminal or network initiates the handover, a network discovery mechanism is required to determine the available base stations for handover. During the network discovery mechanism, mobile terminal scans for the nearby base stations. The network discovery phase is challenging especially when the mobile terminal is moving since there will be many base stations which might be available for handover. Also, the attachment point can change as the handover can be initiated multiple times. When a new network comes in the vicinity of the mobile terminal, the base station announces its presence by sending a message which will identify its presence.
- **Network Selection:** After checking the available base stations, the network controller has to select one base station to handover. There might be more than one base station available for the handover. There are various schemes to select the base station. In the traditional handover scheme, handover is done based on RSSs, where the base station is selected depending upon whether the new base station will provide RSS higher than the current base station.
- **Handover decision:** The handover decision is based on criteria such as available resources. There should be base stations available which can support the current ongoing call. If there are no base stations available, or the available base station cannot provide the quality which the current base station is providing, there is no point in making handover. If the handover is done to such a base station the call might eventually be dropped.
- **Authentication:** Once the decision of the handover is made, and the new base station is selected for the handover, the new base station has to be verified for authentication and authorization. The information of the mobile terminals e.g., the identification number, call type and QoS criteria and billing information are sent to the new base station. The new base station checks and issues an AAA certificate.

2.4 Related Work

The ITU-T definition of NGWN includes the ability to make use of multiple broadband transport technologies and to support seamless mobility. NGWN

must integrate several IP-based access technologies in a seamless way. Many researchers around the world are actively participating to enable and enhance the handover for the NGWN. In (Salsano, Polidoro, Mingardi, Niccolini and Veltri, 2008) the authors provided a survey of the current available handover technologies depending on the requirements of multimedia applications. In addition a SIP-based mobility protocol was proposed by taking into account the requirements of the NGWNs. SIP protocol being the default session establishment protocol for the NGWN, the authors have not achieved much improvement in QoS and handover delay. The authors in (Yan, Ahmet Sekercioglu and Narayanan, 2010) survey currently available handover schemes for NGWN and provided a comprehensive opinion about their suitability to achieve the required QoS in NGWNs. Unfortunately currently proposed Vertical Handover Decision (VHD) algorithms lack a comprehensive consideration of various network parameters. The studies reporting these algorithms lack enough detail for implementation. Research on vertical handover decision algorithms in heterogeneous networks is still a challenging area. The main difficulty is devising an algorithm which is truly useful in a wide range of conditions and user preferences. In (Vegni, Tamea, Inzerilli and Cusani, 2009), the authors compared three handover decision criteria, each of them based on different physical metrics. The authors proposed a smart combination of several handover criteria, which uses not only RSS or SINR parameters, but also a hybrid mixture from different wireless access networks. To improve quality of service for mobile terminals, a data rate gain parameter is introduced that allows an efficient handover to be executed. Data rate gain parameter is not a significant factor and also the handover protocol is limited to three networks which will not add much value.

NGWN mobile terminals are expected to provide multiple network connections such as WiFi, WiMAX, LTE and others. To meet the challenges of the applications it is very important for network providers to dynamically negotiability QoS in the network. One of the initiatives is the bandwidth aggregation scheme. Bandwidth aggregation implies transmission of data from multiple paths. However, the data received at the receiver might be out of order due to various factors such as delay and interference etc.. This further introduces delays to capture all the data frames missing from different paths (Fernandez, Taleb, Guizani and Kato, 2009). LTE-advanced supports carrier aggregation technique to increase spectral efficiency. Carrier aggregation allows deployment bandwidths of up to

100 MHz, enabling peak target data rates in excess of 1 Gb/s in the DL and 500 Mb/s in the UL to be achieved (Ghosh, Ratasuk, Mondal, Mangalvedhe and Thomas, 2010). Another important technique for achieving better spectral efficiency is cognitive radio. Cognitive radio has emerged as another important technique for achieving better spectral efficiency in NGWNs. Learning and adaptation are two significant features of a cognitive radio transceiver. Many Intelligent algorithms are used to learn the surrounding environment, and then utilized by the transceiver to choose the better frequency channel of transmission to achieve the best performance (Niyato and Hossain, 2009).

NGWN devices equipped with capability of accessing multiple networks, face a huge challenge of energy and power management. The applications include high definition voice, video streaming require more power. Another aspect of high power consumption is where the mobile terminal need to scan multiple access networks in various situations. Scanning of these networks consume high power. Hence, it is a challenge for mobile vendors and network operators to have efficient technology to reduce power and energy loss. To minimize and optimize user equipment power consumption, and further to support various services and large amounts of data transmissions, advanced power conservation mechanisms are being developed in IEEE 802.16m and 3GPP. Two advanced power conservation mechanisms, sleep and idle modes, which are enhanced versions of the legacy IEEE 802.16 system's sleep and idle modes, were proposed and adopted in IEEE 802.16m (Kim and Mohanty, 2010). Network operators have realized that the current BSs are not energy efficient which serves many Mobile Terminals. In this regard, the Mobile Virtual Center of Excellence (VCE) Green Radio project was established in 2009 to establish how significant energy savings may be obtained in future wireless systems. The main objective of this project is to create innovative methods for the reduction of total energy to run a radio access network (Han and Harrold, 2011). The authors in (Edler and Lundberg, 2004) take customer perspective, user site, network, climate, and traffic statistics from operator networks to accurately estimate the energy consumption in real network operations. The authors described that energy consumption in the usage phase of its radio access networks is the most critical factor relating to impact on the environment.

Chapter 3

Performance Evaluation of Quality of VoIP in WiMAX and UMTS

The NGWNs are supposed to be a convergence of different radio access technologies, and are taking a tremendous leap, with features like global roaming, anytime anywhere wireless access to the Internet, and automatic handover in heterogeneous networks. For a mobile terminal to use different radio networks, several inter-working architectures between UMTS and WiMAX have been proposed (Nguyen-Vuong, Fiat and Agoulmine, 2006). Even though current mobile terminals are compatible with different radio access technologies, networks like WiMAX and UMTS provide different QoS for real-time applications in the current development. Making a handover from one network to another is difficult due to the complexity of the networks. One major difficulty for call handover from one network to another is the guarantee of end-to-end QoS. To get the best available service for a user, it is necessary to study the QoS model of these networks, i.e. to study their QoS differences and possible ways to resolve the differences between their QoS models. We believe that the classification of different QoS requirements from the real-time multimedia applications will help to choose the best available network without degrading the QoS of the applications. These classifications can be used for implementing the resource scheduling system of UMTS and WiMAX when they are integrated. For example, depending upon the network congestion and available resources, a call can be transferred from

UMTS to WiMAX or vice versa to provide better QoS to the users.

The QoS can be studied between LTE and WiMAX or LTE and UMTS. In this work we study QoS between UMTS and WiMAX. Firstly, LTE is a new technology which is yet to be launched in many countries. Secondly, LTE which is being launched in few countries are being launched only for data support and not for voice. The reason for this is the voice over LTE (VoLTE) is supposed to be ready by 2014 or 2015 except one or two countries which will launch the voice service on LTE in 2014. The only country that has VoLTE is South Korea at the moment. Further, a subscriber on LTE falls back to UMTS/3G for voice service until VoLTE is completely implemented. In addition, VoLTE needs a separate platform called IMS to support voice. Considering all these aspects I chose UMTS and WiMAX which are two deployed networks available in the market.

In this work, we take VoIP as an application scenario to study the differences of QoS between UMTS and WiMAX, in order to investigate how well these two networks cope with real-time multimedia applications. This study will help to identify the strengths and weaknesses of the two networks in terms of QoS and can guide the applications to choose the best available network in a heterogeneous environment. We have designed and implemented WiMAX and UMTS simulation modules in OPNET (OPNET, 2010) and carried out extensive simulations to analyze the MOS, packet end-to-end delay, jitter and packet delay variation for different type of VoIP traffic in these two networks. Our simulation results show that WiMAX has better QoS to support VoIP compared with UMTS.

The rest of the chapter is organized as follows. Section 3.1 briefly describes the background including VoIP, QoS in UMTS and WiMAX. Section 3.2 deals with the simulation setup used in OPNET for both UMTS and WiMAX. Section 3.3 evaluates and analyzes the simulation results of the VoIP application running on UMTS and WiMAX. Section 3.4 deals with related work. Section 3.5 presents the discussion. Finally, in Section 3.6 we present the future work.

3.1 Preliminaries

3.1.1 VoIP

Before the evolution of VoIP technology, telecommunication operated over a PSTN. With the growth of wireless communication, new protocols such as VoIP have emerged. VoIP is a transmission protocol for voice signal over the Internet and is also popularly known as IP telephony. VoIP converts the analog voice signal into digital from the users' equipment and sends over the Internet. When a VoIP call is initiated, the voice signal is first broken down into small packets. These packets are appended with the header, which includes the IP address of the destination. It also contains the routing information. This process is called packetization. The packets are compressed and sent over to the destination address where it is decompressed back. VoIP is mainly used to support multimedia application, such as video conferencing. VoIP can also be used for data applications such as file transfers and chatting.

With the new emerging set of mobile phones, such as iPhones, VoIP has become a de facto standard for voice applications in the Internet. With the telecom industry moving towards the NGWNs which are going to provide high-quality service and higher down-link/up-link speed, VoIP continues to improve its QoS, especially for the long distance calls. This improvement is going to impact businesses like call centers, multinational companies, as well as the normal users to a great extent than ever imagined.

VoIP has become one of the most popular and inexpensive technologies to communicate for short and long distances (Velte and Velte, 2006). VoIP allows communications on existing IP networks without adding extra infrastructure. The transmitted signal is real-time data and addressed to the IP destination. A VoIP call uses both packet-switched and circuit-switched technologies. PSTN just has circuit-switched network through which it is possible to send only voice signal. VoIP uses PSTN for transferring the voice signal and packet-switched network for IP signals. One of the most notable advantages of VoIP is that it can make a long distance call at a cheaper rate than the traditional PSTN (DESANTIS, 2006) (Hallock, 2004) (Khan, 2008).

As seen in Figure 3.1, the VoIP protocol is divided into control plane and data plane (JDSU, 2003) (Chen, Wang, Xuan, Li, Min and Zhao, 2003). The control

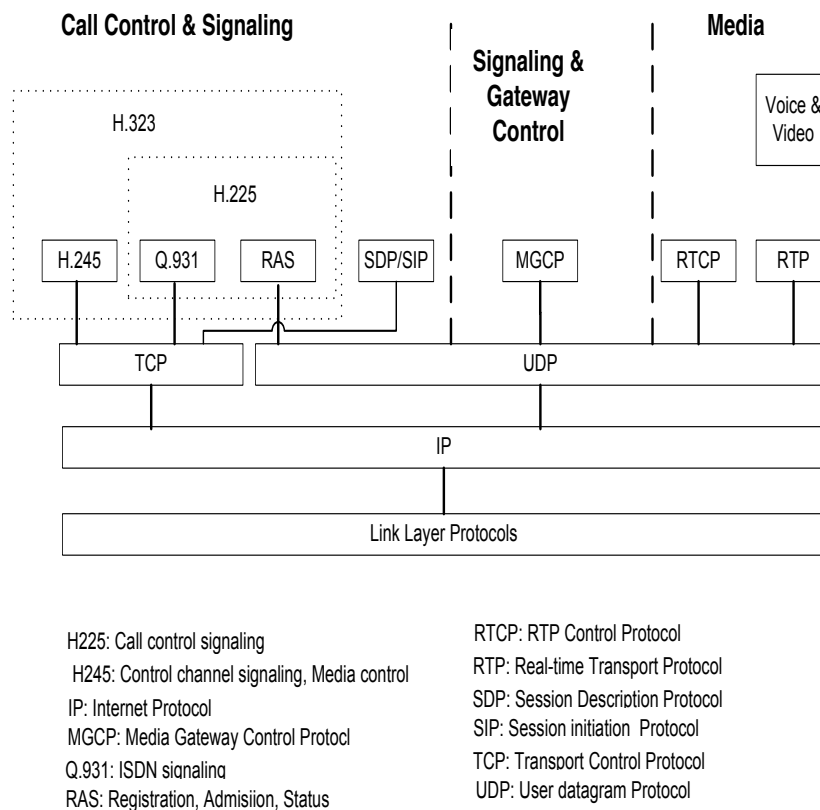


Figure 3.1: Voice over Internet Protocol

plane deals with controlling of the traffic in the network. It connects a user traffic to the network and maintains the connection. The function of control plane includes traffic controlling, traffic monitoring, admission control and resource allocation. Control plane also decides how many resources can be allocated to the multimedia traffic. At the top layer of the control plane, protocols, such as SIP and H.323, initiate the session. As seen in the Figure 3.1, VoIP uses both the Transmission Control Protocol (TCP) and User Data-gram Protocol (UDP). TCP in the control plane sits over the IP layer which is connected to the link layer.

The data plane (voice) of the VoIP protocol is the actual user traffic, which has to be transferred from the source to the destination address. The function of data plane includes QoS management. In VoIP, signaling traffic and voice traffic take different paths. Real-Time Protocol (RTP) supports the voice traffic and the Real-Time Control Protocol (RTCP) is used for the signaling purpose.

VoIP components: The components of VoIP are end user equipment, network components, call processors, gateways and protocols (Institute InfoSec Reading Room, 2004).

- **End user equipment:** The end user equipment is mainly used for accessing the VoIP system in order to communicate with the destination point. The end user equipment can be a mobile phone, computer, laptop or any device that can support the Internet.
- **Network components:** The network components include cabling, routers, switches, firewalls, circuit-switched network and packet-switched network. It uses PSTN lines for circuit-switching.
- **Call processor:** The call processor functions can include phone number to IP translation process, setting up the call, monitoring the call, user authorization, signal coordination, and control bandwidth utilization.
- **Gateways:** The gateways can be categorized into three main types: Signaling Gateways (SG), Media Gateways (MG) and Media Controllers. Signaling gateways control the signal traffic between an IP network and a circuit-switched network like UMTS. The media gateways manage the media signals between them. Media Gateway Controllers (MGC) manages traffic between SGs and MGs.
- **Protocols:** There are many protocols for VoIP. The two most widely used are H.323 and Session Initiation Protocol (SIP).

H.323: One of the popular protocol for a voice/video call on an IP network is H.323. This protocol has wide support among telecom providers and manufacturers all over the world. Specification for H.323 is defined by ITU. This protocol is meant to provide a gateway for telephony devices into the PSTN. H.323 is mostly used in small LANs. One major drawback of H.323 protocol is its lack of scalability.

SIP: VoIP community uses SIP protocol for signaling. This protocol is designed by Cisco systems and is a dominant open standard VoIP protocol. SIP is a signaling protocol specified by the Internet Engineering Task Force (IETF). It is used to set up and tear down two-way communication sessions. There is a SIP server which routes/manages the calls. SIP translates a user name to the current network address, manages the call admission, call

dropping, call transferring mechanisms and allows for changing the features of a session.

3.1.2 QoS in UMTS

UMTS is proposed to converge packet-switched and circuit-switched networks. Its IP Multimedia Subsystem (IMS) is used for multimedia communications. IMS was originally defined by the 3GPP for the next generation mobile networking applications and uses SIP as the signaling protocol. There are some smart phone applications which basically change the natural behavior of the smart phone and let VoIP applications use the 3G network instead of the default WiFi. This application, VoIP over 3G, works with the alternative installer Fring (Shinder, 2007).

So far, four service types have been proposed and incorporated into the QoS model of UMTS:

- Conversational class - Conversational class of QoS is used for the real-time applications which require remarkably low end-to-end delay. This class of QoS requires guaranteed minimum bit rate from the network. Conversational class of QoS is used for voice/video telephony and two-way communication applications requiring low end-to-end delay and low jitter. In case of the UMTS network, the conversational QoS is widely used for the speech service. The speech service can be normal voice call, which is sent through circuit-switched network or an IP-based VoIP call sent through the packet-switched network. This class of QoS maintains the time between the information of the data stream. They have strict delay requirements and are mainly used to carry the real-time traffic.
- Streaming class - This class of QoS maintains the time between the information of the data stream. They have delay and jitter requirements, but not as strict as conversational class. This is because streaming class of service deals with asymmetric data, which is steady and continuous, thus more tolerant to delay and jitter than the conversational class of QoS. This class of QoS does not require guaranteed minimum bit rate from the network. Streaming and conversational class of QoS are both used for real-time application. The conversational class has strict requirement for the delay than the stream-

ing class of QoS. This class of QoS is used for the Internet-based streaming application, mostly multimedia streaming. It allows the live compressed real-time data, or pre-recorded data to be delivered in real-time.

- Interactive class - This class of QoS in UMTS is used for the online applications with low packet loss rate and two-way data transfer. Examples of interactive class are the Internet-based applications, such as web browsing, File Transfer Protocol (FTP) and telnet. Interactive class of QoS is delay tolerant than conversational and streaming class. They have liberal requirement for the delay. This traffic is characterized by request-response communication pattern.
- Background class - Background class is also known as best effort class of QoS. It deals with traffic which does not require immediate delivery, but has to reach within tolerable delay. This QoS class is suitable for the non-real-time applications. It is used for the Internet based applications, such as web browsing, emails, downloading and Short Message Service (SMS). These applications are not time sensitive and do not require high priority. It can be processed and delivered at the background. They are used for one way communication traffic and can have higher packet loss rate. In this QoS class, there is no stringent delay requirement for the packets to reach the destination. They have a re-transmission process in which the lost packets can be re-transmitted.

In UMTS, VoIP traffic is routed directly from the GGSN to the VoIP server. Voice signal is sampled, digitized, encoded, and decoded in mobile terminal. SIP is used as a signaling protocol in 3GPP. The voice and video stream packets are carried over a protocol called RTP which is encapsulated in UDP. The protocol stack for UMTS has RTP/UDP/IP headers that have 40 bytes (IPv4) in total (60 bytes in IPv6) followed by voice data as payload. RTP has a 12-byte header; UDP has a 8-byte header, and IP has 20-byte header.

3.1.3 QoS in WiMAX

QoS in WiMAX

So far, five service types have been proposed and incorporated into the QoS model of WiMAX (Tranzeno, 2010):

- Unsolicited Grant Service (UGS) - Supports real-time data streams for delay constraint traffic which require optimal throughput. UGS supports jitter tolerant, maximum latency tolerant (5-40 ms latency over the air and 100 ms latency over an IP backbone) and maximum sustained rate applications. UGS supports application with Constant Bit Rate (CBR) service, such as VoIP for which achieving low latency is extremely critical. In WiMAX, UGS flows are buffered separately from the other service classes, such as nrtPS and Best Effort (BE), so they get higher priority over other trivial applications, such as SMS. During the upstream, the system uses UGS to bypass the normal request-grant mechanism for upstream traffic by allowing the base station to give automatic grant to a UGS flow.
- Real-time Polling Service (rtPS) - Supports real-time data streams. It is used for real-time services, such as streaming video that generates the data packets of variable sizes with variable bit rates, a guaranteed minimum rate and a guaranteed delay. The rtPS has more request overhead than UGS, but it supports variable grant sizes for data transport efficiency. Unlike UGS, there is a polling overhead which can sometime reach up to 60 percent. The rtPS supports periodic, high priority, maximum latency tolerance, maximum reserved rate and maximum sustained rate applications. A drawback of this QoS type is that it has a significant impact on the overall throughput.
- Non-real-time Polling Service (nrtPS) - Supports delay tolerant data with variable packet sizes. The nrtPS service class supports non-real-time services that require variable size data packets, and a minimum data rate with higher latency, such as FTP. This is done by using unicast polls on a regular basis, which ensures that the service flow receives requests even during network congestion. Priority is given to UGS and rtPS applications over nrtPS.
- Best Effort (BE) - Supports data streams where no minimum data rate is required, and packets are handled based on available bandwidth. Unicast polling requests are not guaranteed in this case, requiring contention requests to be used. BE packets may take long time to transmit during network congestion.
- Extended real-time Polling Service (ertPS): This type of QoS is used for scheduling algorithms for VoIP service with variable data rates and silence suppression. This service class has been newly introduced to support real-

time service flows that generate variable sized data packets on a periodic basis with minimum reserved rate, maximum sustained rate, maximum latency tolerance, jitter tolerance and traffic priority. The ertPS service class enables silence suppression mechanism and makes better use of header compression. If the conversational class uses voice traffic with talk and silence spurt detection, then it can be mapped to the ertPS class of WiMAX networks. VoIP is an example of ertPS class application.

Table 3.1: Mobile WiMAX Applications and Quality of Service

QoS Category	Application
UGS (Unsolicited Grant Service)	VoIP
rtPS (Real-Time Polling Service)	MPEG
nrtPS (Non-Real-Time Polling Service)	FTP TFTP HTTP
BE (Best-Effort Service)	Email Data Transfer Web Browsing
ErtPS (Extended Real-Time Polling Service)	Voice with Activity Detection, VoIP

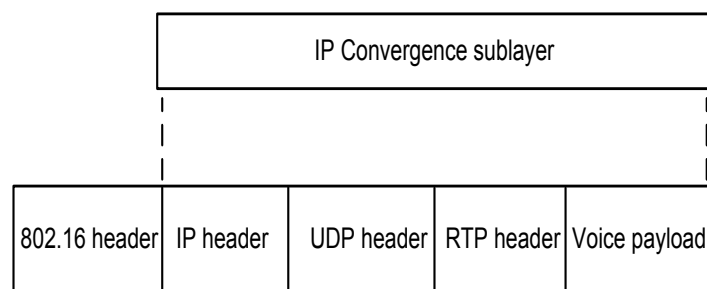


Figure 3.2: Frame Structure of WiMAX

A VoIP packet, as shown in Figure 3.2, consists of RTP/UDP/IP headers. Since WiMAX supports variable-sized frames, the header size of RTP can be between 12 and 72 bytes and the header size of UDP is 40 bytes. The total length of RTP/UDP/IP headers ranges between 60 and 120 bytes. Given the larger size of

the RTP/UDP/IP headers, the proportion of actual voice payload per packet is only 33. Both WiMAX and UMTS have advantages and disadvantages compared to each other. WiMAX is the first truly open mobile standard (IEEE 802.16e) governed by the IEEE's fair licensing practices and open to participation. This is in fact revolutionary since 3GPP and 3GPP2 are consortium and do not allow open participation. This open process should lead to greater innovation and hence a better performance when moving forward and can potentially reduce intellectual property licensing fees because it provides a quicker improvement of the technology compared to existing mobile technologies. WiMAX is also the first major mobile standard to offer all-IP network. UMTS will get there in subsequent releases, but it still employs a complicated and ultimately expensive core network (Visant, 2008). Table 3.1 shows the mobile WiMAX applications and their corresponding QoS categories.

3.2 Simulation Setup

One of the objective for the work in Chapter 3 is to show which technology is better for users to select applications they want to use. Hence the performance comparison between the two technologies is done. The difference of the channel capacity impacts the queuing delay, however this is not the only factor. There are other factors, such as network switching delay, QoS and support for the application which significantly impact the delay. As to be shown later in Section 4.1, the QoS is related to delay and packet loss. UMTS provides best effort QoS for most of the applications, whereas WiMAX has dedicated QoS model for different applications. In our simulation, in order to have fair comparison, we have used best effort service for both WiMAX and UMTS.

The main objective of our study is to evaluate the performance of VoIP traffic in UMTS and WiMAX network. To evaluate the performance of WiMAX and UMTS for VoIP traffic, we have designed and implemented WiMAX and UMTS simulation modules in the OPNET network simulator (OPNET, 2010) based on OPNET's discrete event simulation model library. In the OPNET, simulation is done over the packet-level. We selected the VoIP application for this study, as VoIP is an IP-based application and will be most commonly used in the NGWN for communication. We investigate the QoS parameters, such as MOS, end to end

QoS, packet delay variation and jitter for both WiMAX and UMTS networks.

3.2.1 WiMAX Simulation Module

To measure the QoS parameters for a VoIP call on WiMAX, we use a setup as shown in Figure 3.3. Each user supports VoIP application. A user can be associated with a single call at a time. Each user terminal is configured with the VoIP connection. We use G.711 encoder scheme as it supports Pulse Code Modulation (PCM). We developed WiMAX model based on IEEE 802.16 standard by using components from the OPNET library. Each user is associated with one base station which is connected to ASN-GW. ASN-GW is connected to the Internet. Each user is either at a fixed position or moving. The movement of the moving user is random and is based on the random walk model.

As illustrated by Figure 3.3, the WiMAX simulation module is composed of mobile terminals denoted by UEs, Base Station (BS), and Access Service Networks-Gateway (ASN-GW). The BS provides air interface to the UEs for VoIP calls and is also responsible for tunnel establishment and radio resource management. The BS is connected to the ASN-GW which is responsible for connection management, location management, radio resource management, admission control, caching of subscriber profiles and AAA client functionality. When a call is initiated the AAA server provides the following functions:

- (a) Authentication: Confirmation that a user requesting a network service is entitled to do so. It involves the presentation of an identity and credentials, such as the user name, password, and/or digital certificate.
- (b) Authorization: Granting of specific types of service based on the subscription of the users.
- (c) Accounting: Tracking the consumption of network resource by the users.

When a mobile terminal (e.g. UE0) makes a VoIP call to another user (e.g. UE1), signaling protocol such as SIP is used to setup the route for the transmission over the IP network. In our simulation we use the SIP protocol. The call is authenticated at the AAA server, and special services will be granted based on the subscription of the users. A channel is then opened on which the actual media will travel using UDP for transport. The Gateway protocols like the Media Gateway Control Protocol are used to establish control and status in the media

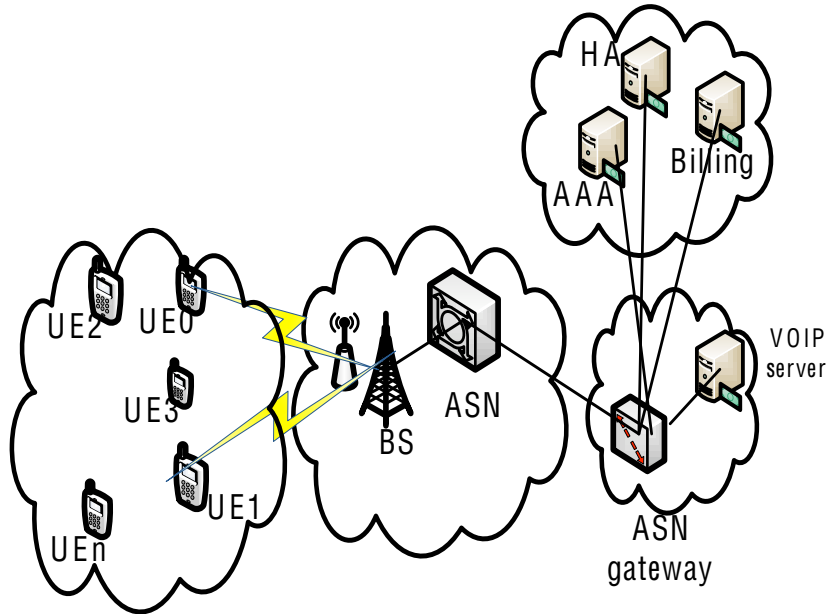


Figure 3.3: VoIP in WiMAX

and signaling gateways. Routing and Transport Protocols (RTP) are used once the route is established for the transport of the data stream.

3.2.2 UMTS Simulation Module

To measure the QoS parameters for a VoIP call on UMTS, we use the setup as shown in the Figure 3.4. We configure the VoIP applications on all the mobile terminals so that each user can make and accept a VoIP call. Each mobile terminal transmits VoIP packets to another user. At a given time, a user can accept only one call. When a call is initiated from UE0 to UE1, the call is processed if UE1 is not busy and the network has sufficient capacity. Once the call is processed, we use G.711 encoder scheme, which supports PCM. We developed an UMTS model based on the 3GPP standard by using the components from the OPNET library. Each user is connected to one Node-B, which acts as base station in the UMTS network. The application config is configured to the VoIP application. When a VoIP call is made it goes through the RNC, SGSN and GGSN. GGSN is connected to the Internet. Each user is either at fixed position or moving. The movement of the moving user is random and is based on the random walk model.

As illustrated by Figure 3.4, the designed UMTS simulation module consists of

the Core Network (CN), UMTS Terrestrial Radio Access Network (UTRAN) and UE. CN provides routing, switching, and network management functions. The radio interface for UE is provided by UTRAN which comprises of Radio Network Controller (RNC) and Node B (base station) (Holma and Toskala, 2006b). A UE can be a mobile handset, a laptop, a desktop or any device that can provide access to the network. The accounting information (time usage, type of service) and subscriber status is forwarded to the Home Subscriber Server (HSS) and billing. AAA performs authorization, authentication and accounting of the call.

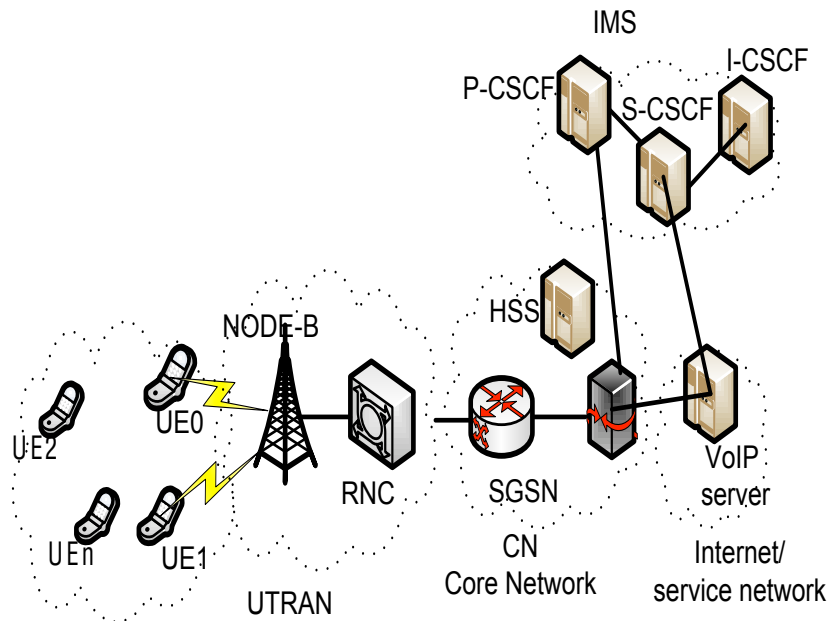


Figure 3.4: VoIP in UMTS

Figure 3.4 shows an overview of VoIP in UMTS. Similar to WiMAX, when a mobile user (e.g. UE0) makes a VoIP call to another user (e.g. UE1), the voice packets are carried from the UEs to RNC through Node-B over a protocol called Real-time Transport Protocol (RTP) which is encapsulated in User Datagram Protocol (UDP). First the Radio Resource Control (RRC) connection is established over the channel. Then, Radio Network Controller (RNC) sets up a point-to-point radio connection as well as the signaling connection to the network before sending acknowledgment back to the UE. The accounting information (time usage, type of service) and subscriber status is forwarded to the home AAA server for authorization, authentication and accounting of the call.

After that, the UE will start the attach process. Then, the Packet Data Protocol (PDP) context will be set up. The PDP context contains mapping and routing information for packet transmission between the UE, SGSN and the gateway GSN (GGSN). In UMTS, VoIP traffic is routed directly from the Gateway to the VoIP server. Voice signal is sampled, digitized, encoded, and decoded in UE. SIP is used as a signaling protocol in 3GPP. PCM quality voice has been generated over IP. The BE type of service and the weighted round-robin queuing have been selected.

3.2.3 Simulation Configuration

To make fair comparisons, we use the same VoIP configurations for UMTS and WiMAX networks. Table II shows the setup. The service flow is designed to support best effort type of service with variable size data packets as VoIP with silence suppression, since both UMTS and WiMAX support it. There are different codec, such as G.711, G.721 and G.722 etc. We use the G.711 encoder scheme which supports PCM. We use one voice frame per packet as it increases the call handling capacity of the network compared with using two or three voice frames per packet. The algorithmic compression delay for G.711 is 0.02 and decompression delay is 0.02 seconds.

Pulse Code Modulation (PCM) quality voice has been generated over IP. The Best Effort (BE) type of service with bronze service class and initial Quadrature Phase Shift Keying (QPSK) modulation with 12 initial coding rate is used for the setup. The average Service Data Unit (SDU) is 120 bytes and the buffer has the size of 64KB.

3.2.4 Performance Metrics

In our simulations, we use the following four metrics to evaluate the performance of WiMAX and UMTS in terms of end-to-end QoS for VoIP. We used IEEE E-Model to calculate the QoS parameters (Sophia, 1999) and (ITU, 2000).

- **Mean Opinion Score (MOS)**

MOS is based on various factors that affect the voice quality and is calculated by the subjective test. In VoIP applications, the call quality is traditionally mea-

Table 3.2: WiMAX and UMTS Network Parameters

Attribute	Values
Silence Length (seconds)	200ms
Talk Spurt Length (seconds)	1.366 sec
Symbolic Destination Name	Voice Destination
Encoder Scheme	G.711
Voice Frames per Packet	1
Type of Service	Best Effort (0)
RSVP Parameters	None
Traffic Mix	All Discrete
Signaling	None
Compression Delay (seconds)	0.02
Decompression Delay (seconds)	0.02

sured from a user's perception using MOS in a range varying from 1 to 5 (Balan, Eggert, Niccolini and Brunner, 2007) (ITU-T, 1996): "1" indicates extremely poor quality and difficult for communication, "5" indicates exceptionally good quality and perfect for communication. The MOS value does not have to be a whole number. MOS value is very commonly used to measure the perceived audio quality of voice signals in VoIP. It indicates whether the call quality is poor or good. MOS value above 3.2 is considered acceptable for communication.

The measurement of the MOS is immensely useful as it helps to evaluate the call quality and takes the respective measures such as handing over of the call to another base station. The factors that affect MOS are:

- Resources
- Jitter
- Codec
- End to end delay
- Packet loss.

MOS is a direct function of delay and packet loss. It increases with the decrease of delay and packet loss. The higher the MOS value, the better the call quality.

If there are sufficient resources available in the network, it reduces the delay and improves the MOS value. The codec used significantly affect the MOS quality. G.711, also called as PCM, is a very commonly used codec. G.711 gives the best quality as it does not require compression and decompression which makes it less susceptible to the latency and packet loss. It reduces the bandwidth usage by using lossless data compression techniques. There are other codecs, such as G729, G722 and G723 etc., which have higher packetization delay incurring higher total delay and decreasing the MOS. The best codecs provide efficient compression and decompression and can save the resources by conserving the bandwidth. Whereas, if the codecs performance is poor, it leads to more bandwidth utilization and also affects the quality of the voice signal.

In our simulation, we compute MOS from R-score as in (ITU, 2000):

$$M = 1 + 0.035R + 7 * 10^{-6}R(R - 60)(100 - R) \quad (3.1)$$

where $R = 100 - I_s - I_e - I_d + A$. I_s is the effect of impairments that occur with the voice signal; I_e is the impairments caused by different types of losses occurred due to codecs and network, and I_d represents the impairment caused by delay particularly mouth-to-ear delay. A is the advantage factor. Using the default setting for I_s and A , R can be simplified to $R = 94.2 - I_e - I_d$.

The R-score ranges from 0 to 100. R-score of more than 70 usually means a VoIP stream of acceptable quality. The relation between the R-score and the MOS rating is given by (Passito, Mota, Aguiar, Carvalho, Moura, Briglia and Bids, 2005),

$$MOS = \begin{cases} 1, & \text{For } R < 6.5, \\ M \text{ given by Equation (3.1),} & \text{For } 6.5 \leq R \leq 100, \\ 4.5, & \text{For } R > 100 \end{cases} \quad (3.2)$$

- **Packet end-to-end delay**

End-to-End delay is the time taken for a packet to be transmitted across the network from source to destination. End to end delay directly impacts the call

quality and affects the system performance. The higher the end to end delay, the poorer the quality of the call. Higher delay might cause packet loss. End to end delay is caused from several factors, such as network condition, various components, type of application, available resources and routers. It is categorized into

- Processing delay: This is caused by hardware components such as codec. Encoding and decoding are the example of the processing delay, which can cause significant impact on processing the signal.
- Propagation delay: This is the delay due to network conditions. Network conditions significantly affect the signal propagation especially of the wireless signal.

The routing delay and channel delay also significantly impact the end to end quality of signal.

Different applications have different requirements for the end to end delay. Multimedia applications such as video conferencing, streaming, VoIP and on-line gaming have particularly stringent requirements for the end to end delay. To provide short end to end delay in UMTS and WiMAX, QoS is categorized based on the applications. The critical applications get higher priority over the trivial applications. End to end delay can be minimized by having efficient scheduling schemes.

In wireless communication, radio interference causes a significant delay and affects the system performance. There is also a significant amount of delay during the handover process. When the call quality of the ongoing session is going down, the mobile terminal initiates the handover process. It requests the base station to transfer the call from the current base station to the new base station, which can provide better call quality. This involves a tedious and time consuming process which can affect the end to end delay. There are various handover schemes proposed to carry out the call transfer efficiently without affecting the call quality.

The total voice packet delay is calculated as:

$$D_{e2e} = D_n + D_e + D_d + D_c + D_{dc} \quad (3.3)$$

where D_n , D_e , D_d , D_c and D_{dc} represent the network, encoding, decoding, compression and decompression delay, respectively.

- **Jitter**

Jitter in the VoIP is the variation of the delay over different packets. It is measured as the time difference between the consecutive packets and appears to be a shaky pulse causing a deviation to the signal. The deviation can be measured with respect to the change in the frequency, amplitude or the width of the signal from the ideal. Jitter significantly impacts the quality of the application and is included as a performance parameter due to its importance at the transport layer in packetized data systems (Wang and Prasad, 2005). There are various sources of jitter, such as limited buffer size, network congestion and delay due to longer routes for the packets which affect the communication link quality. The power supply noise, electromagnetic interference and other noises in the communication also cause the jitter in the signal.

The most common cause of jitter in wireless mobile communication is limited buffer size. With advancement in technology, the data rate of mobile communication is increasing at a faster pace. But the mobile terminal has limited buffer capacity, and they can only handle limited packets at a time. This leads to packets arriving at different latencies, which causes the jitter. When there is a constant latency in the packet arrival there is no jitter as there is no variation. Sometimes the latency of the packets may be longer and buffer might get full, this might lead to dropping few packets causing the packet loss.

The jitter has a significant impact on the quality of the application as it causes fluctuations in the signal. Different applications have different thresholds for the allowable jitter. The quality of the application significantly deteriorates if the jitter value is above a tolerable limit especially a real-time applications like VoIP are significantly affected with large jitter. Jitter can be reduced by increasing the buffer size or by using de-jitter buffer. There are various schemes proposed to find the optimal size of the de-jitter buffer for various applications. There are schemes proposed to create a buffer that re-sequences packets as they come in and passes them on to the application.

In OPNET, jitter is computed as the signed maximum difference in one way delay of the packets over a particular time interval. Let $t(i)$ and $t'(i)$ be the time transmitted at the transmitter and the time received at the receiver, respectively. Jitter is calculated as follows:

$$jitter = \max_{i=2}^n ([t'(i) - t'(i-1)] - [t(i) - t(i-1)]) \quad (3.4)$$

- **Packet Delay Variation (PDV)**

Packet delay variation is the difference in the end to end delay between the selected packets. Usually consecutive packets are selected for the calculation of PDV. The PDV is one way delay metric. The variance of packet delay can be used to estimate the variance of the inter-packet delay variation (Duffield and Lo Presti, 2000). PDV is used for the sizing of play-out buffers for applications such as voice or video play-out that require the regular packet transmission. The maximum delay variation plays a crucial role and is used to size play-out buffers (Demichelis and Chimento, 2002). If PDV is high, it can be a critical problem for real-time applications such as video conferencing and VoIP, hence the WiMAX and UMTS networks have a quality of service model where the packets are prioritized depending upon whether the application is real-time or non-real-time. PDV is significantly affected by the network load.

PDV in OPNET is defined as the variance of the packet delay, which is computed as follows:

$$PDV = \frac{\sum_{i=1}^n ([t'(i) - t(i)] - u)^2}{n} \quad (3.5)$$

where u is the average delay of the n selected packets.

3.3 Simulation Results and Analysis

In this section, we compare the performance of VoIP in WiMAX and UMTS networks through extensive simulations. To effectively analyze the performance, we measure the four metrics presented in Section 3.2.4 over a set of simulations with different number of homogeneous mobile users. All the mobile users in the same simulation use the same configuration.

3.3.1 MOS

Quality of a call is determined by the MOS value. The higher the MOS value the better is the call quality. Figure 5.7 plots the average MOS with different

number of homogeneous VoIP connections.

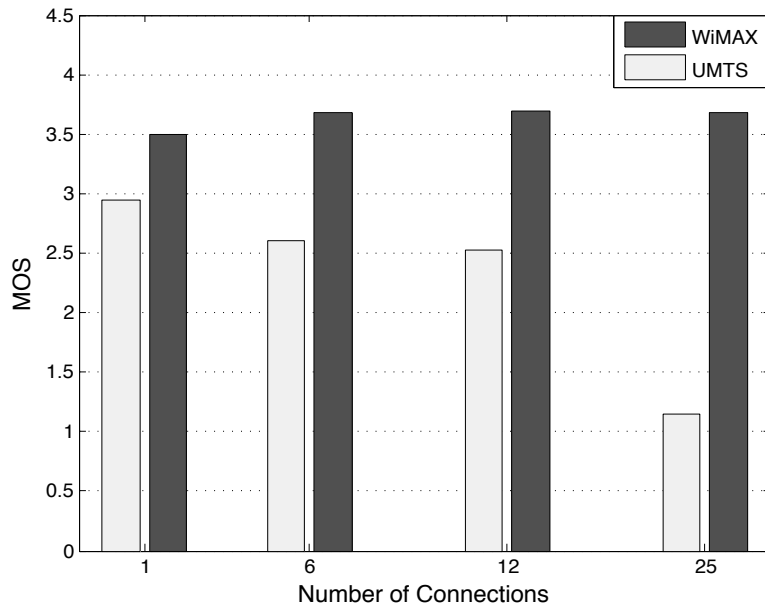


Figure 3.5: Tendencies of MOS with regards to number of connections

A major observation is that the average MOS decreases with the increase in the number of connections in UMTS, whereas the average MOS remains roughly steady in WiMAX irrespective of the number of VoIP connections. With 25 connections, the MOS in WiMAX is almost 3 times larger than that in UMTS. As can be seen from Eqn (3.1), the MOS value depends on the R-score which is a function of the packet loss and delay. The higher the packet loss rate and delay, the lower the MOS value. This indicates that WiMAX has smaller delay and packet loss rates as compared with the UMTS network and can support VoIP applications better than UMTS.

The WiMAX network is a packet switched network and can support IP-based applications like VoIP very well. Since the WiMAX is an all-IP network there is no circuit switching involved. All the calls go through the packet switched network. On the other hand, UMTS is a circuit-switched as well as packet switched network. When a call is made in the UMTS network, it has to go through the selection process where, depending upon whether it is IP-based call or a voice call, switching is done. This switching process in the UMTS network leads to an extra delay in the call processing, which leads to an increase in the overall end-to-end delay of the system. Delay and packet loss are very sensitive parameters for real-time applications such as VoIP as they significantly affect the audio quality

of the signal. From the results in Figure 3.5 we can see that the VoIP application can be run on UMTS as well as a WiMAX network, but the WiMAX network can provide better quality. Another reason for large delays in the UMTS network is UMTS is the fact that UMTS is a very big network comprised of small cells which are interconnected to provide larger coverage. When a call is made in a UMTS network it has to travel to many components such as base station, RNC, SGSN and GGSN before it gets to the destination causing additional delay in the network. The WiMAX network on the other hand is smaller and has fewer components in the network. When a call is initiated in the WiMAX network, it travels through fewer components, reducing the travel time of the packets.

This suggests that, compared with UMTS, WiMAX has fewer congestions, less traffic burst and better bandwidth allocation strategies, and thus low packet loss rate. It is clear that WiMAX can provide better voice quality than UMTS, especially in scenarios with a large number of VoIP connections. There are several other factors like QoS model, resource management and scheduling which affects the system delay and packet loss impacting the MOS value.

3.3.2 Packet End-to-End Delay

Packet end-to-end delay is one of the most important performance metric in VoIP. Packet end-to-end delay directly affects the QoS of a call. We measure and evaluate one way delay in UMTS and WiMAX networks with different scenarios. Figure 3.6 and Figure 3.7 show the average packet end-to-end delay in UMTS and WiMAX, respectively. We plot the measurement from the time (i.e. 160th second) when the communications become stable as it takes some time to set up the VoIP connections. As can be seen from the figures, the average delay in WiMAX is much more steady than that in UMTS. With 25 homogeneous connections, WiMAX has an average packet end-to-end delay of 0.09 second, which is less than 50% of the average delay in UMTS. When the number of connections is increased from 1 to 6 and 6 to 12, the average end-to-end delay is increased by 33% and 25%, respectively in UMTS, whereas in WiMAX it is increased only by 14.2% and 12.5%.

From Figure 3.6, it can be seen that there is an increase in the delay after 250 secs. This is because the UMTS network supports best effort class of QoS. Since there is no separate queue for different applications based on the priority, all the

applications pass through the same queue. Hence, if the number of users increases the delay also increases with respect to the time. In case of WiMAX, applications are classified depending on the QoS requirements and are passed through different queues. Due to this WiMAX has better delay handling capability if the number of users increases.

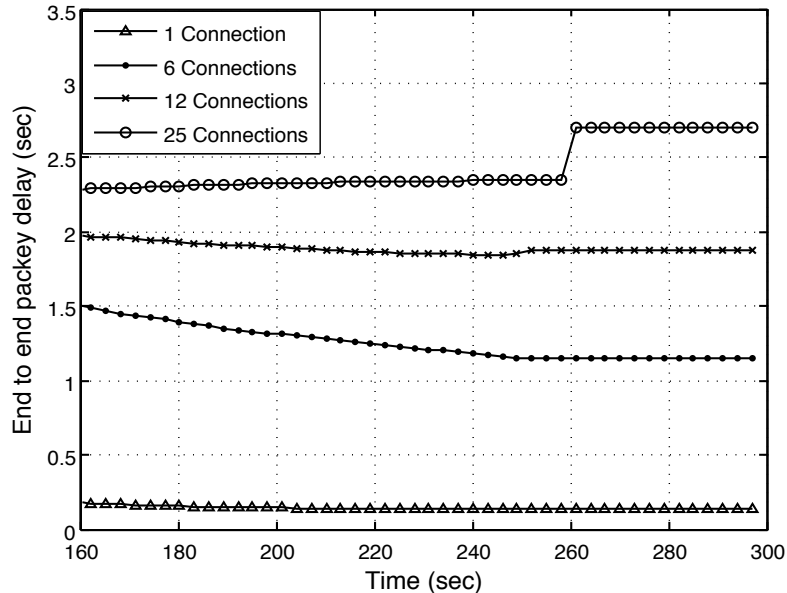


Figure 3.6: Packet End-to-End Delay in UMTS

The simulation results indicate that WiMAX can provide better VoIP services in terms of end-to-end packet delay. The reason is that WiMAX is an all-IP network, whereas UMTS is still a combination of circuit and packet-switched technologies. A VoIP call in UMTS has to go through a selection procedure to choose the circuit-switched network or the packet switched network, which takes a considerable amount of time contributing to the end-to-end delay of the network. The above results of QoS will immensely help in selecting the proper radio access interface for multiple-interface mobile terminal, which will become common very soon. The delay significantly impacts the scheduling properties in the network.

Another reason for higher delay in the UMTS network is that, although UMTS has QoS model for which it assigns QoS class according to priority of the application, the current UMTS network only supports best effort class of service. Due to this all applications are treated same and are passed through the same queue which results in the delay as the number of users increase. The WiMAX network

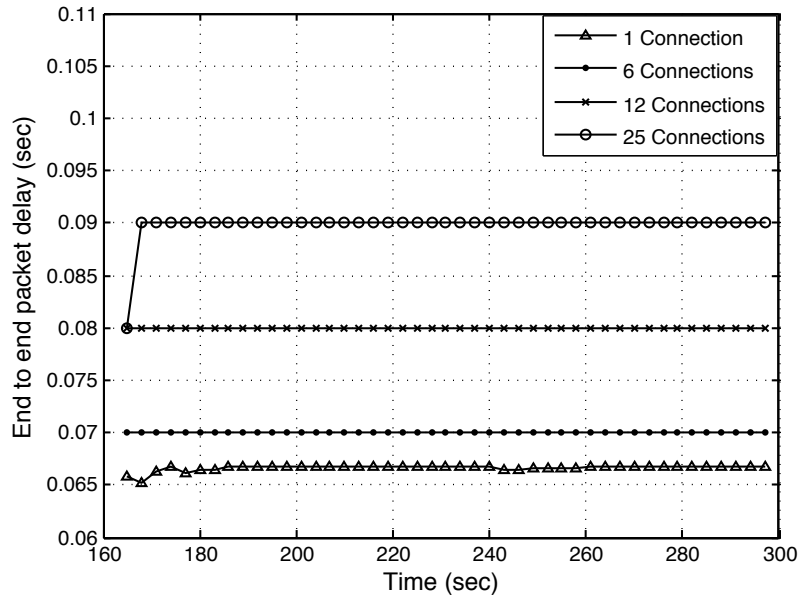


Figure 3.7: Packet End-to-End Delay WiMAX

on the other hand has well defined QoS model. Packets are categorized depending upon whether it is very urgent, real-time, non-real-time or best effort service. Once the categorization is done the packets are sent through the separate queues. Hence the application like VoIP, which is delay sensitive, is assigned UGS class and sent through the respective queue. The QoS model in the WiMAX network significantly impacts on the quality of the call.

The delay in the network also depends on the codec used, encoding and decoding schemes. In our simulation configuration, we use the same codec, encoding and decodings scheme for both UMTS and WiMAX.

3.3.3 Jitter

We compare the jitter value of UMTS and WiMAX for the VoIP application. We use different scenarios, where we increase the number of users. We can see that the jitter value increases when the mobile terminals increases. When a call is initiated packets are transferred in a queue. When there are fewer users in the network the queue is not congested and all the packets sent are delivered without dropping or delaying. When the number of users increases the queue length increases, which causes the variation of the packet arrival causing the jitter. Thus, if the real-time traffic in the network increases, jitter increases. The

non-real-time applications such as e-mails are less susceptible to jitter as they are not time critical.

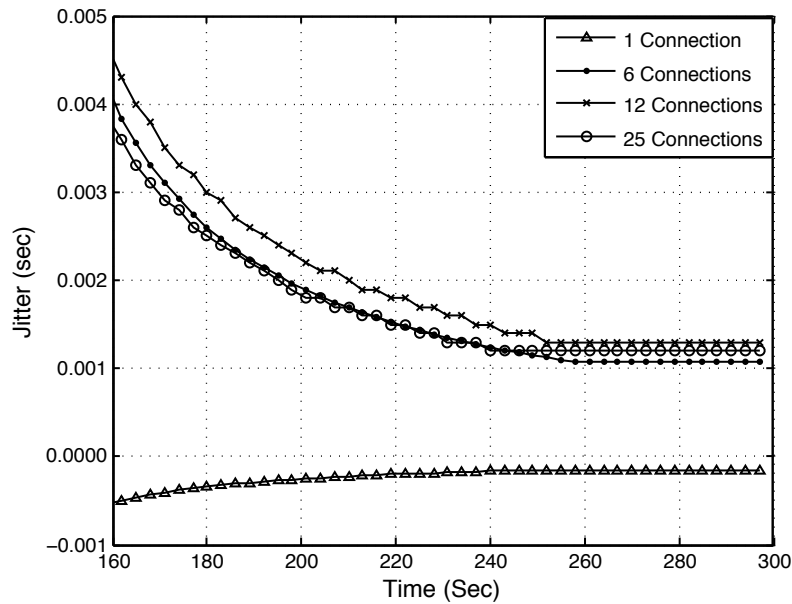


Figure 3.8: Jitter in UMTS

According to Eqn (3.4), the jitter value can be negative which means that the time difference between the packets at the destination is less than that at the source. Figure 3.8 and Figure 3.9 plot the jitter in UMTS and WiMAX, respectively. It can be seen that UMTS has a large range of jitter variation, ranging from -0.0005 to 0.0045, and takes longer time to converge to the stable stage. For WiMAX, it has a narrow range between 0.0000 to 0.0003, accounting for only 6% of that for UMTS. Moreover, it has a fast convergence to the stable state. This phenomenon can be explained as follows: as the number of users increases in UMTS, the congestion in the system also increases due to the slow packet scheduling. The multimedia sessions, such as streaming will end up in more increased time. This will cause delay in transferring packets at the receiving terminal, thereby leading to a poor quality audio for a user. The two figures show a maximum jitter value of 0.0045 for UMTS and 0.0003 for WiMAX, respectively. The jitter factor is also a reason for a decreased MOS value for the UMTS. From Figure 3.8, it can be seen that the UMTS has higher jitter at the start of the call. This is because of the circuit-switching involved at the beginning of the call. In UMTS when a call is initiated it goes through the selection process where, depending upon whether it is IP-based call or a voice call, switching is done. This caused additional jitter

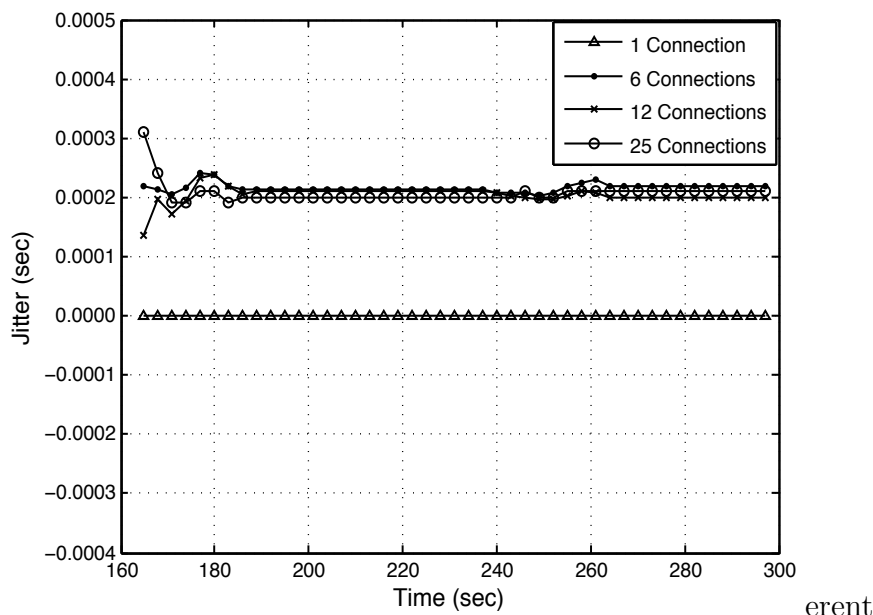


Figure 3.9: Jitter in WiMAX

at the start of the call.

Delay indirectly causes the jitter in the systems. UMTS networks are prone to more delay and packet loss for real-time application like VoIP as they need extra circuit-switching time. Buffer size has a significant impact on the jitter. Results show, WiMAX has a better buffer management system than the UMTS network. An increase in the number of users in the WiMAX network does not result in an increase in the jitter value. Figure 3.9 shows the WiMAX network maintains the jitter value even if the number of users are increased. In the case of UMTS network, the jitter value is acceptable and within the limit, but as the number of users increases the jitter value increases. The jitter value is initially higher as the UMTS network involves switching process but over a period of time it stabilizes. In the case of the WiMAX network as there is no circuit switching process involved the jitter value is stable over the full call duration.

The jitter performance metrics are much broader than transfer of packets. Jitter describes the variation in the data stream rate or carrier phase noise. In the IP network the packet inter-arrival rate is variable. Hence the size of the jitter buffer is adjusted to allow predetermined variation. The jitter and total delay are not the same, although having plenty of jitter in a packet network can increase the amount of total delay in the network. This is because the more jitter, the larger the jitter buffer needs to be to compensate for the unpredictable nature of the

packet network. It can be seen in Figure 3.11 there is a packet delay variation in the WiMAX network in order to maintain the jitter.

3.3.4 Packet Delay Variation

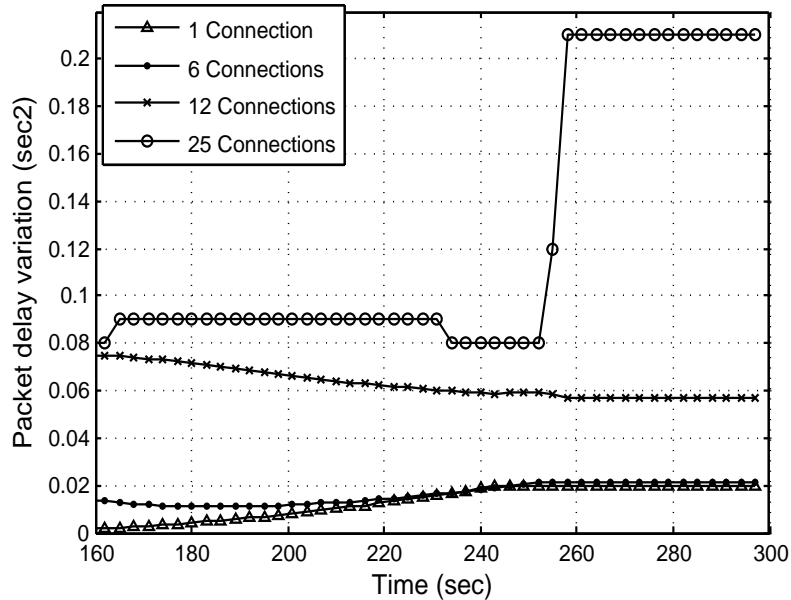


Figure 3.10: Packet Delay Variation in UMTS

Packet delay variation plays a crucial role in the network performance degradation and affects a user perceptual quality. Higher packet delay variation results in the congestion of the packets which can result in the network overhead. This is because if there is a variation in the time at which the packets are sent, they will not be received by the receiver at the intended time and have to be re-sent. This will cause higher load in the network. Low packet delay variation is important for applications requiring timely delivery of packets, e.g. VoIP, video etc. The maximum delay variation is useful for determining the optimal buffer sizes for such applications.

Figure 3.10 and 3.11 show that, WiMAX has a smaller delay variation of 0.00015 which can be tolerated because of buffering and jitter compensation within the voice decoder, thereby providing a stable QoS for the service. UMTS on the other hand has a larger delay variation of 0.21 seconds, and this results in disturbed QoS particularly in streaming services. Results show that, WiMAX networks are more stable and reliable than UMTS networks. In the real-time application such

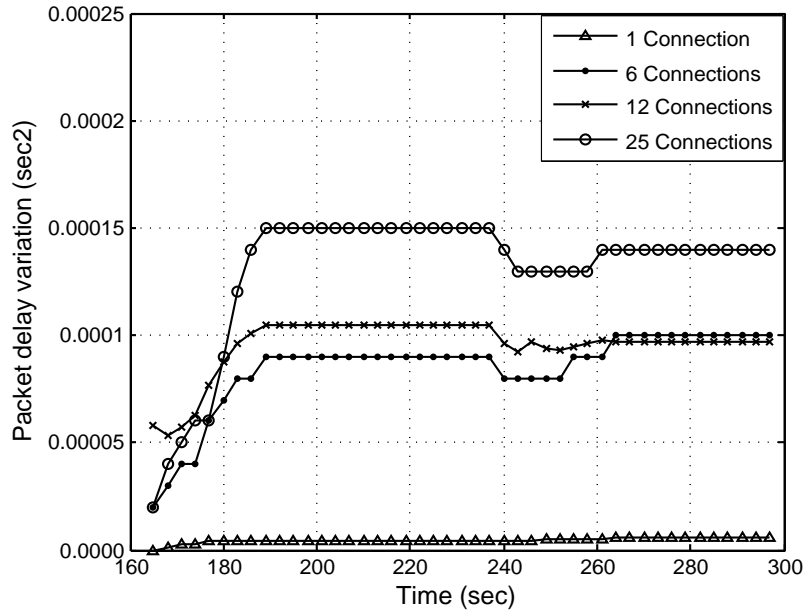


Figure 3.11: Packet Delay Variation in WiMAX

as Video conferencing, PDV plays a critical role as it directly impacts the quality of the call. UMTS is likely to have higher PDV than WiMAX due to the reasons mentioned in the above sections.

3.4 Related Work

Most of the existing work was done to evaluate the performance of VoIP in either WiMAX or UMTS. The authors in (Kim, 2006; Lunden, Aijanen, Aho and Ristaniemi, 2008) evaluated the capacity of VoIP services on High-Speed Down-link Packet Access (HSDPA), in which frame-bundling is incorporated to reduce the effect of relatively large headers in the IP/UDP/RTP layers. This work concluded that the capacity of VoIP service on HSDPA is attractive for transmission of voice. The authors in (Garg and Yu, 2000) discussed the quality of service requirement for traffic class in the 3G UMTS network. The authors suggested that, we need to realize QoS requirement in MAC/LAC sublayer along with the restriction imposed by W-CDMA air interface. The authors in (Soldani, 2005) determined the role and importance of some of the key aspects of QoS planning, provisioning, monitoring and optimization for UMTS within the framework of the 3GPP. The authors explained the differences between Quality of Service Experience (QoE) and QoS concept in 3GPP. Through the simulation results the

authors proposed that, if QoS mechanism is configured properly network can considerably reduce the bandwidth.

The work in (Lee, Kwon, Cho, Lim and Chang, 2006) analyzed the efficiency of resource utilization and VoIP capacity in IEEE 802.16e, and showed that UGS and rtPS algorithms have problems in supporting VoIP, such as waste of up-link resources in the UGS algorithm and the additional access delay and MAC overhead due to bandwidth request process in the rtPS algorithm. It was stated that the ertPS algorithm can support 21% and 35% more voice users compared with the UGS and rtPS algorithms. The authors in (Adhicandra, 2010) examined QoS deployment over cellular WiMAX networks, and compared the performance of VoIP application using two different QoS configurations (UGS and ertPS). Their results showed that ertPS had advantages in scenarios with delay-sensitive traffic. The work in (Scalabrino, De Pellegrini, Chlamtac, Ghittino and Pera, 2006) reported on the measurements on a real WiMAX network through synthetic VoIP traffic generation. Although some work had been devoted to understand different QoS models of a particular network with respect to VoIP, there is not much work on comparing the performance of QoS of VoIP traffic in different networks. Different with the existing work, we focus on evaluating QoS parameters for VoIP on two popular and widely deploying networks, UMTS and WiMAX, which will be inter-operating with each other in the near future.

The authors in (Li, Qin, Low and Gwee, 2007b) presented an overview of the IEEE 802.16 MAC protocol and deals with the issues associated with scheduling and QoS provisioning. The authors also discussed the main features of the newly standardized mobile WiMAX, IEEE 802.16e. The authors discussed QoS provisioning and mobile WiMAX specification. In (Belghith and Nuaymi, 2008), authors investigated the behavior of WiMAX scheduling algorithms and focused on the rtPS class. The authors pointed out the problem that can arise due to rtPS service class and proposed a solution to provide more spectrum efficient scheduling. Paper (Belghith, Nuaymi and Maillé, 2008) investigated pricing schemes for a WiMAX system with different classes of QoS and focused on BE class pricing. The authors proposed two pricing schemes for BE: Fixed Symbol Price Model and Variable Symbol Price Model (FSPM and VSPM, respectively). VSPM was based on auction. Through simulation authors showed that, FSPM can provide a higher revenue than VSPM. FSPM does not take into account the satisfaction of the BE users, and the operator has to change the symbol price at each network

change.

In (Kim, Cai, Na and Choi, 2008), authors conducted multiple experiments in various environments to analyze the performance of commercial mobile WiMAX network. The authors conducted an experiment where, multiple users with UDP packets fully utilize wireless links. The authors analyzed few performance parameters like the good-put and round trip delay. The measured performances of mobile WiMAX are compared with the HSDPA which is an enhanced 3G network. The parameters were analyzed for the indoor and outdoor environments. It was found that, WiMAX has better performance over HSDPA. In the indoor environment, performance of both the networks was reduced but still WiMAX had better performance over HSDPA. The experiments were conducted with the users moving around the city. It was observed that WiMAX was able to provide the seamless service to the users. The authors found that WiMAX system performs a role for mobile broadband wireless access in Seoul, Korea despite some known problems which need to be fixed. The authors in (Sengupta, Chatterjee and Ganguly, 2008), exploited the features offered by the WiMAX at MAC layer for the construction and of MAC transmission protocol data units, which helped to support greater numbers of VoIP calls. The authors proposed the combination of techniques to enhance the performance and support multiple VoIP streams.

The authors in paper (Arjona, Westphal, Yla-Jaaski and Kristensson, 2008) analyzed the quality of VoIP on HSDPA network. With the simulation, the authors showed HSDPA can significantly reduce the user-to-user voice delay, but this is only satisfactory for few devices and the overall end-to-end QoS experience is substantially worse than circuit-switched solutions and not acceptable. The authors proposed that the current disadvantage of the HSDPA with VoIP application can be reduced, and it can potentially decrease the jitter buffer size reducing the terminal processing delay. The authors in (Braga, Rodrigues and Cavalcanti, 2006) studied the packet scheduling and evaluated the packet scheduling algorithm for guaranteeing the QoS for the VoIP application. The authors considered a scenario where VoIP and web browsing services competed for the same resources. The authors showed that packet scheduling algorithm is able to perform QoS differentiation and manage the capacity resources well. Thus, the performance of VoIP application can be improved by using proper packet scheduling algorithm. The above studies showed that, WiMAX and UMTS networks have their own

advantages and disadvantages. By integrating these two networks, user will be able to get most of the advantages of both the network, and will be able to get better end to end QoS. There are inter-networking architectures proposed. The authors in (Khan, Ismail, and Dimyati, 2010) proposed an inter-working architecture between UMTS and WiMAX to maintain the connectivity of an on-going call. The authors proposed to keep the Internet Protocol (IP) address to be static and mobile terminal with multiple transceivers with the ability of making handover decisions. Through the simulation, using OPNET the authors showed that this approach can eliminate the packet loss due to handover. In (Nguyen-Vuong, Fiat, and Agoulmine, 2006), the authors proposed a UMTS-WiMAX inter-working architecture based on 3GPP standards and proposed a handover procedure which promises a low packet loss and low interruption time during the switching of the communication.

3.5 DISCUSSION

In last few years communication technology has undergone major changes. Old copper wires are getting replaced with fiber optics, new networks are getting introduced which support faster data rates and wider coverage. In the next few years the focus is on the NGWN, which is supposed to be the convergence of the all networks providing faster speed, better quality, seamless mobility and wider coverage. NGWN is an integration of the different networks such as UMTS, WiMAX and WLAN etc.

Moving to the NGWN is not an easy task. Currently these networks operate independently and have different requirements for data rates, coverage, bandwidth and application support. These networks support real-time multimedia applications but they have different architectures and system to process the call. Many potential problems have to be resolved to make switching between networks smoother. For example, UMTS and WiMAX have different QoS models. It is challenging to map the types of QoS between the two networks for an ongoing session. The mobile terminal should be auto-configured to a suitable network according to the QoS requirements of the application in use. This requires an extensive study of how different applications behave on different networks, in order to select the best available network. The chosen network, on the other

hand, will have a problem to match the QoS parameters to those of the previous network for an ongoing session. In this regard, this work will help in application classification and QoS restructuring for an ongoing session handed over between different networks.

In future, when the UMTS and WiMAX will allow the users to select any of the available networks, the classification of QoS makes it easier for the network operators to allow a user to switch to the network which is best suited for real-time applications.

Chapter 4

MOS-based Handover Protocol for Next Generation Wireless Networks

The envisaged NGWNs will integrate a number of different networks, such as UMTS and WiMAX to provide a comprehensive and secure all-IP based services to mobile terminals. Future mobile terminals will be equipped with multiple network interface cards, which enable the mobile users to connect to different networks and access any service anywhere and anytime. However, heterogeneous networks are different in data rates, traffic classes and call admission mechanisms etc. How to seamlessly transfer user service between networks of the same type or between different networks is a well-known handover issue, and has become one of the major issue in developing and deploying the NGWNs.

With the rapid growth of wireless packet-switched networks, the Internet has become more popular than the PSTN in terms of cost for both users and service providers. This has led to enormous growth of real-time applications based on VoIP, enabling the mobile terminals to make calls through Internet anywhere and anytime with better communication quality and less cost than PSTN. With a growing number of moving users, it has become a necessity to guarantee the QoS for applications that demand more bandwidth, better network connectivity and seamless handover. Moreover, wireless networks are susceptible to delay, packet loss and poor call quality due to the low Signal to Interference and Noise Ratio (SINR). An efficient handover management scheme should be designed to

achieve better call quality in NGWNs.

Mobility management plays an important role in maintaining end-to-end QoS. Mobility management in NGWN network basically consists of location management and handover management. The location management deals with finding the location of the network in which mobile terminal is attached to, change of attachment from the current network and delivering the data to the mobile terminals. The handover management involves managing the active service continuity on the mobile terminal with minimum delay and no loss of packets. Seamless handover will maintain the same QoS of the currently running service whereas basic handover may or may not consider or provide the same set of QoS.

The main objective of handover is to maintain the quality of a call, in other words to maintain the delay and packet loss below threshold limit. It can occur due to several reasons, such as quality of service going down, network capacity not sufficient or network might not support the application. The handover process is carried out in milli-seconds and it goes unnoticed by the users. Handover might require to transfer a call from one radio access technology to another. Different radio access technologies have differences in the way the call is handled. Nevertheless, they have differences in type of application support, data rate and QoS which make the handover process more complex. NGWN is a heterogeneous network and has been designed to support the seamless mobility of the users within any network.

In the existing handover schemes, a handover is generally triggered by either the detection of degradation in Received Signal Strength (RSS) or using other metrics, such as measurement from network load, power consumption, user preference and available bandwidth. The traditional handover protocols based on RSS or cost functions (Yan, Ahmet Sekercioglu, and Narayanan, 2010) have flaws and are not competitive enough to achieve satisfactory QoS. The mobile terminal has to scan continuously for the current and available networks signal strength. This scanning procedure utilizes wireless resources and also encounters with the wireless channel access delay. Mobile terminals with continuous scanning also consume more battery power, thereby resulting in energy inefficiency. Another problem is the signal fading which will give rise to the ping-pong effect, resulting in unnecessary handover (Chang and Chen, 2008). In the RSS-based handover the signal strength depends upon the the distance of the mobile terminal from

the base station. The authors in (Grønsund, Grøndalen, Breivik and Engelstad, 2007) have given relationship between RSS and distance which is shown in Figure 4.1. As shown in the figure RSS value decreases as the mobile terminal moves away from the base station.

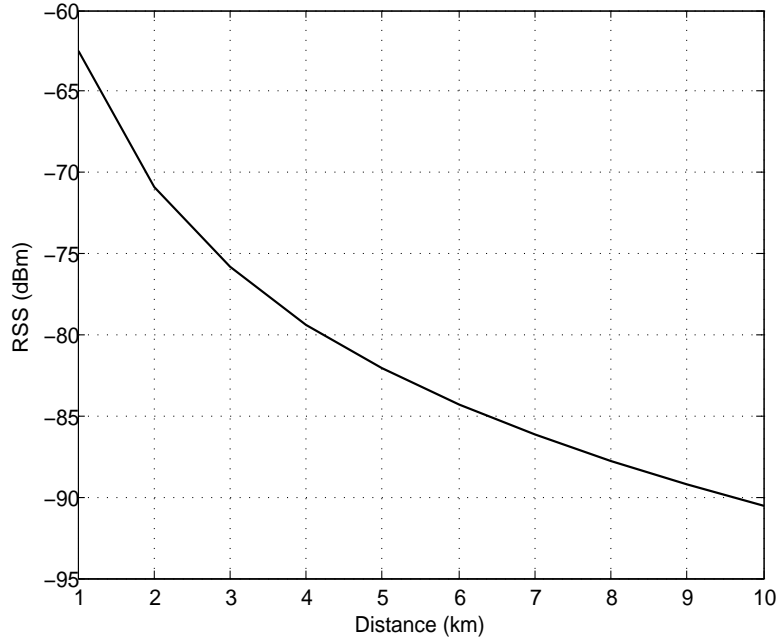


Figure 4.1: RSS vs Distance

In this chapter, we investigate the following problem: *for an urban area deployed with both UMTS and WiMAX base stations, when a mobile terminal experiences degrading call quality from its current connection, how can we choose an optimal base station for the mobile terminal to handover in terms of maximizing the call quality measured by Mean Opinion Score (MOS)?* We model this problem as an optimization problem by considering the available bandwidth at the base stations, the communication delay and loss, and the MOS values. A centralized algorithm is designed to compute the optimal base station for handover. To enable the handover between base stations both in the same network and in different networks, we design a handover protocol compliant with the recently proposed IEEE 802.21 standard, which is also called as the Media Independent Handover (MIH) (IEEE802.21, 2008). The standard defines a media-independent handover framework that can significantly reduce the complexity for handover between heterogeneous network technologies. We have done extensive numerical simulations to evaluate our MOS-based handover protocol. Simulation results

show that our scheme can provide much better performance than the traditional RSS-based handover schemes.

The rest of the chapter is organized as follows. Section 4.1 discusses the related work. Section 4.2 briefly discusses the E-model, which provides the expected voice quality prediction, as perceived by a mobile terminal for an end-to-end connection. Section 4.3 presents the problem formulation. Section 4.4 gives the optimal base station selection algorithm. Section 4.5 describes the handover protocol design. Section 4.6 describes the simulation setup. Section 4.7 presents the numerical results and finally, in Section 4.8 we present the discussion.

4.1 Related Work

Most of the existing work on handover in UMTS, WLAN and WiMAX is based on bandwidth (Oliveira, Kim and Suda, 1998), SINR (Yang, Gondal, Qiu and Dooley, 2007) or RSS (Kunarak and Suleesathira, 2010) (Chang, Chen, Hsieh and Liang, 2009). Yang et al in (Yang, Wu and ROC, 2007) proposed that, when roaming from WiMAX networks to Wi-Fi networks, it is reasonable to initialize handover to Wi-Fi when Wi-Fi is available because Wi-Fi networks can provide high bandwidth and lower cost. However, they do not consider the handover probability. When a user is moving and handover is required, Wi-Fi network can be very small and a user might need to handover again, thus increasing the handover probability and affecting QoS. The authors in (Saboji and Akki, 2011) proposed a mobile agent based bandwidth estimation mechanism which is applied during the handover of connection-less and connection-oriented services. This algorithm is based on agent technology. The agents are created to measure parameters related to bandwidth availability. There are several disadvantages of this scheme: 1. An agent platform is required in the mobile terminal, wireless access point and control node; 2. Identifying and exporting agent state information are critical problems; 3. Additional services are required for authorization and portability of the agent; 4. Additional encoding techniques are required for managing data, state and code in secrete state; 5. A large number of agent migrations results in significant overhead in terms of both latency and required data to be transported across the network. In (Liang, 2002), the authors presented a soft handover scheme for non-uniformly-loaded mobile cellular networks. Band-

width utilizations of two mobile terminals are combined with signal strength of two base stations to make a 4-D soft handover possibility surface. Hard decision of handover is finally determined based on the category of multimedia and QoS. In (Yang, Gondal, Qiu, and Dooley, 2007), a handover algorithm is proposed to use the received SINR and bandwidth from various access networks as the handover criteria. Through simulation authors showed the SINR based vertical handover algorithm is able to consistently offer the end user with maximum available throughput during vertical handover. However, selecting a candidate network with significantly higher bandwidth, is not always beneficial to the serving network. There are different environmental and networking factors which cause variation in SINR. This results in increase in the handover probability and might cause unnecessary handover.

The work in (Lee and Kim, 2002) is based on forced termination of calls due to handover failure. The dropping of a handover call is generally considered more serious than blocking of a new call. Therefore, a certain amount of bandwidth (also called guard channels) is exclusively reserved for handover. This amount of bandwidth can be either fixed or adaptively controlled with respect to the current traffic load. RSS and bandwidth are important factors but there are several other factors which might degrade the quality of voice signal. A user might be just standing beside the base station and there might be sufficient bandwidth available, but the network to which a user is attached might not support VoIP call well, or there might be other networks available which might provide better quality. In our scheme, we not only consider bandwidth but also take into consideration the quality of service parameters.

There are location-based handover algorithms proposed, which consider the mobile terminal location as a criteria for the handover. The authors in (Inzerilli, Vegni, Neri and Cusani, 2008) presented soft mobile-controlled vertical handover between wireless systems based on the location. In this scheme the authors introduced a preliminary handover initiation phase which is triggered on the basis of the mobile terminal location. The goodput is then estimated and handover decision is taken. This approach optimizes the good-put and limit ping-pong effect. The authors in (Lin, Juang and Lin, 2005) presented a handover algorithm based on the location and velocity of the mobile terminal to suppress the ping-pong effect in cellular systems. An approach with reasonable errors was used

to estimate the location and velocity to identify a correlation among shadowing effects. The proposed scheme was applied to a real GSM system in urban Taipei city. The computational complexity of the algorithm is low and no database or lookup table is required.

The work in (Ma, Yu, Leung and Randhawa, 2004) is based on a method to facilitate seamless vertical handover between wide-area cellular data networks such as UMTS and WLANs using the Stream Control Transmission Protocol (SCTP). In a UMTS/WLAN overlay architecture, the multi-homing capability and dynamic address configuration extension of SCTP are applied. This decreased the handover delay and improved throughput performance. This scheme was a network-independent scheme, as it does not require the addition of components such as home/foreign agents or a SIP server to existing networks. The authors in (Choi and Cho, 2005) introduce a new concept, takeover, which enabled the network to process requests of the mobile terminal. Through the simulation results the authors showed takeover-based vertical handover scheme achieved faster and seamless handover compared with conventional handover schemes without the takeover.

The authors in (Liao, Tie and Du, 2006) presented a vertical handover decision algorithm based on the fuzzy control theory. The authors considered multiple criteria, such as Power Level, Cost and Bandwidth for making the handover. A membership function is established after which membership degrees of corresponding factors is determined. These are then processed by the Weight Vector. Lastly, vertical handover decision is made based on the Fuzzy vertical handover decision vector. However, considering multiple factors and fuzzy algorithm, the complexity and processing time increases. This incurs delay in making the handover. The authors in (Dai *et al.*, 2008) proposed an approach for vertical handover between WiMAX and WiFi networks. The proposed algorithm combines data rate and channel occupancy in order to fairly balance users among the two networks. The authors simulated this algorithm in an urban environment. However, in this work the authors did not consider the QoS for handover. The authors in (Desset, Ahmed and Dejonghe, 2009) optimized handover decisions between WLAN (802.11) and WiMAX (802.16e) standards, for both up-link and down-link data transmission. The authors derived power and performance models for both standards and evaluated the opportunity of handover. Factors such as channel fading fluctuations, extraction of MAC-level behavior, packet error rates and

overall power consumption from the wireless platform were taken into consideration. The authors then designed a handover controller that selects the network with the lowest expected power for the required rate. However, this mechanism is very complicated since it is based on multiple factors. In addition, this algorithm is based on regular scanning and involves high energy consumption.

4.2 E-model

In wireless communications, MOS is a widely used metric to measure the perceived quality of a voice call. The quality of service factor is estimated from an E-Model provided by the ITU-T. This factor is called as R-score. The E-Model (Sophia, 1999) and (ITU, 2000) has been used to define the R-score which is a function of voice impairment factors in the network. The R-score is given by

$$R = R_o - I_s - I_e - I_d + A \quad (4.1)$$

where R_o is the effect of noise, I_s is the impairment caused by signal-to-noise, I_e is the equipment impairment factor which can be caused by codec due to its lower data rate, I_d is the impairment caused by sum of all the delays and A is the advantage factor and takes care for the above impairment under various user condition. R_o , A and I_s are considered to be fixed. I_e and I_d are the main factors affecting the quality of voice signal.

4.2.1 Effect of Delay in Packet Network

The factor I_d is called as mouth-to-ear delay, which means the sum of all delays the packet encounters while traveling from mouth (i.e., source) to the ear (i.e., destination). This delay is comprised of encoding, decoding, buffer, de-jitter and queuing delay. The I_d value that is less than 177.3ms does not affect the voice quality and is acceptable for the communication. According to the E-model, the I_d for VoIP steam is given by

$$I_d = 0.024 * d + 0.11(d - 177.3)H(d - 177.3) \quad (4.2)$$

where $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x \geq 0$. d is the total mouth-to-ear delay. The total delay d is composed of three components codec delay, play-out delay and network delay.

$$d = d_{codec} + d_{payout} + d_{network} \quad (4.3)$$

where d_{codec} is the delay caused by coder and encoder. d_{payout} is the play-out delay which is caused due to buffering and $d_{network}$ is the delay in the network. The d_{codec} delay varies from codec to codec and is mainly caused due to the packetization. Table 4.1 show the delay value for the different codecs. As seen from the table G.711 has the least value of delay.

Table 4.1: Delay Values for Codec

Codec	Delay
G. 729	25ms
G.723	67.5ms
G.711	0.25ms

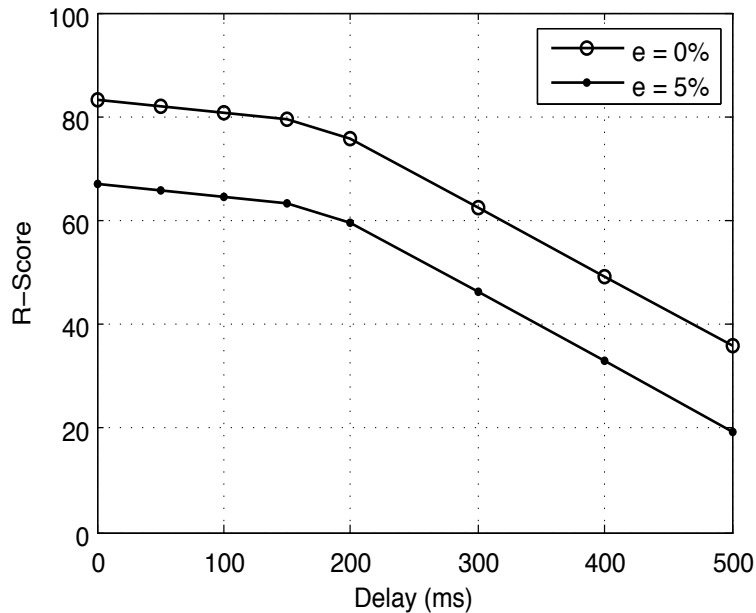


Figure 4.2: R-score vs Delay

Figure 4.2 shows the behavior of delay with respect to the R-score. The packet loss rate is kept constant. As shown in Figure 4.2 the delay increases with the R-

score decreasing. The R-score value above 80 indicates high voice quality. When the delay is low the R-score is high. When the delay value is above 173ms there is sudden decrease in the R-score value.

4.2.2 Effect of Packet Loss Rate in Packet Network

The I_e is represented by the packet loss rate and is caused by network and receiver play-out loss. In the E-model, I_e is computed by

$$I_e = \gamma_1 + \gamma_2 * \ln(1 + \gamma_3 * p) \quad (4.4)$$

where p is the network and play-out buffer losses and $\gamma_1, \gamma_2, \gamma_3$ are constants and are known parameters for given codecs. As seen from Equation (4.4) the quality of the call depends on the codec used. Different codecs have different I_e values. Table 4.2 show the values for packet loss for the different codecs (ITU, 2000).

Table 4.2: Packet Loss Values for Codec

Codec	γ_1	γ_2	γ_3
G.729	0	45	18
G.711	0	30	15

The network and play-out buffer losses p in Equations (4.4) is given by

$$p = p_{network} + (1 - p_{network}) * p_{playout} \quad (4.5)$$

where $p_{network}$ is the packet loss rate in the network and $p_{playout}$ is the packets loss rate during the playout of voice signal.

4.3 Problem Formulation

Consider an urban area where a UMTS network and a WiMAX network coexist, as shown in Figure 4.3. We use the tightly coupled architecture (Xu, Zhang and Zhou, 2007a), in which a single Radio Network Controller (RNC) maintains the network information such as the available capacity at each base station and

the quality of each wireless connection. This can be achieved by requesting each base station to periodically update the resource usage and the quality of currently served applications. All handovers occurred in this area are managed and optimized at the RNC.

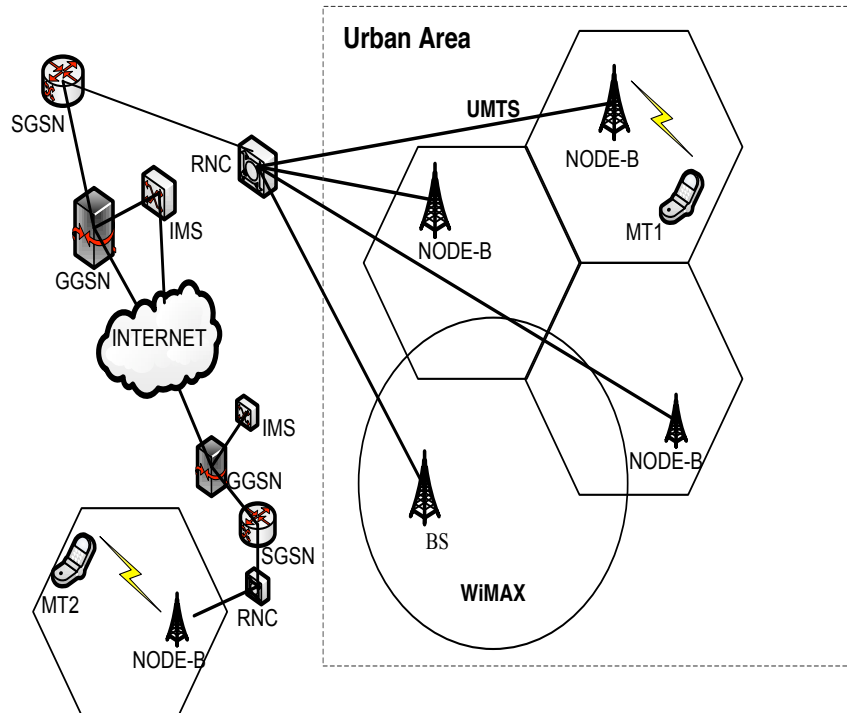


Figure 4.3: Tightly Coupled Architecture for an Integrated Scenario

As shown in Figure 4.3, the connection setup for communication between the mobile terminals $MT1$ and $MT2$ can be divided into two parts: connection between the base stations, and connection between the mobile terminal and base stations. The connection between the base stations goes through the Internet using the wired medium, and voice data is transmitted using the VoIP protocol. Since the Internet commonly has large communication bandwidth, it can provide relatively stable communication quality, thus having little impact on the voice quality. However, the communication between mobile terminal and base stations is wireless. Since wireless channels are prone to errors due to noise interference as well as the movement of the mobile terminal, the quality of the wireless channel generally dominates the quality of the VoIP call. In this study, we focus on the communication between the mobile terminals and the base stations.

Let $B = \{b_1, b_2, \dots, b_n\}$ be the set of UMTS and WiMAX base stations deployed

in urban area, and $M = \{m_1, m_2, \dots, m_n\}$ be the set of mobile terminal. Given a base station b_i and a mobile terminal m_j , let $p_{i,j}$ and $d_{i,j}$ denotes the packet loss rate and the average packet delivery delay between b_i and m_j , respectively. If mobile terminal m_j is not in the coverage range of base station b_i , $p_{i,j} = 1$ and $d_{i,j} = \infty$. Consider a mobile terminal m_i which is experiencing poor call quality, our objective is to design an efficient solution to select the best base station for m_i to handover, by which the call quality is maximized.

In this work MOS will be used as the main metric for handover optimization. It is one of the major metrics for evaluating the quality of a VoIP call. Let $M_{i,j}$ denote the MOS value for the VoIP call between base station b_i and mobile terminal m_j . In (Ding and Goubran, 2003), MOS is computed as follows:

$$M_{i,j} = 1 + 0.035R + 7 * 10^{-6}R(R - 60)(100 - R) \quad (4.6)$$

where $R = 94.2 - I_e - I_d$, I_e is modeled as follows:

$$I_e = \gamma_1 + \gamma_2 * \ln(1 + \gamma_3 * p_{i,j}) \quad (4.7)$$

where $p_{i,j}$ denotes the packet loss rate between base station b_i and mobile terminal m_j .

The I_d for a VoIP steam is given by

$$I_d = 0.024 * d_{i,j} + 0.11(d_{i,j} - 177.3)H(d_{i,j} - 177.3) \quad (4.8)$$

where $d_{i,j}$ denotes the average packet delivery delay between base station b_i and mobile terminal m_j .

Figure 4.4 shows the behavior of delay with respect to the MOS. The packet loss rate is kept constant. As shown in the figure the MOS value decreases with the delay value increasing.

The relation between the R-score and the MOS rating is given by (Passito, Mota, Aguiar, Carvalho, Moura, Briglia, and Bids, 2005),

$$MOS = \begin{cases} 1, & \text{For } R < 6.5, \\ M_{i,j} \text{ given by Equation (4.6),} & \text{For } 6.5 \leq R \leq 100, \\ 4.5, & \text{For } R > 100 \end{cases} \quad (4.9)$$

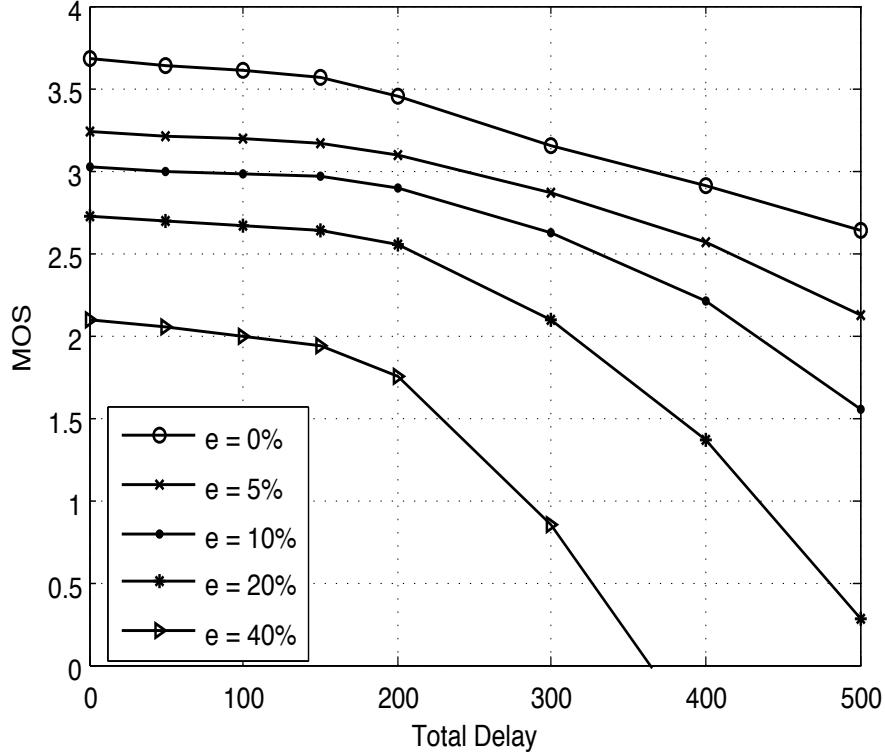


Figure 4.4: MOS vs Delay

By Equations (4.6), (4.8) and (4.7), it can be seen that $M_{i,j}$ can be expressed as a function of packet loss $p_{i,j}$ and delay $d_{i,j}$, i.e., $M_{i,j} = f(p_{i,j}, d_{i,j})$.

Suppose that mobile terminal m_j is currently making a VoIP call demanding bandwidth of \bar{c} and experiences poor call quality, our goal is to choose the best base station that meets the bandwidth requirement for m_j to handover in terms of maximizing $M_{i,j}$. The base station selection problem can be formulated as the following optimization problem:

$$\begin{aligned}
 & \text{maximize} && M_{i,j} = \max_{b_i \in B} f(p_{i,j}, d_{i,j}) \\
 & \text{s.t.} && c_i \geq \bar{c};
 \end{aligned} \tag{4.10}$$

where c_i represents the available bandwidth capacity at the base station b_i .

4.4 Optimal Base Station Selection

In this section, we present the solution for choosing the optimal base station that maximizes the MOS value, assuming that the RNC has the knowledge of the

bandwidth capacity of each base station, and the delay and packet loss rate for each wireless link between mobile terminal and base stations. In VoIP applications, the call quality is traditionally measured from a users' perception using MOS in a range varying from 1 (bad) to 5 (excellent). Details on how to obtain these parameters will be described in next section.

From Equation (4.9), it can be seen that MOS monotonously increases with the increase of R when $0 < R < 100$. By Equations (4.6), (4.7) and (4.8), it is easy to prove that the MOS value monotonously increases with the decrease of packet loss rate and packet delay. Then we have the following observation.

Observation: Given a mobile terminal m_j and two base stations $b_i : (p_{i,j}, d_{i,j})$ and $b_k : (p_{k,j}, d_{k,j})$. If $p_{i,j} \leq p_{k,j}$ and $d_{i,j} \leq d_{k,j}$, we have $M_{i,j} \geq M_{k,j}$.

The above observation enables to quickly drop unsuitable candidates during base station selection process. Let $b_k : (p_{k,j}, d_{k,j})$ be the current severing base station for m_j . The base station selection procedure works as follows: we initially use $p_{k,j}$ and $d_{k,j}$ as the benchmark for the base station selection. Given a base station $b_i : (p_{i,j}, d_{i,j})$ in B ,

1. If $p_{i,j} \geq p_{k,j}$ & $d_{i,j} \geq d_{k,j}$, b_i can not provide better call quality than the current serving base station b_k according to the observation.
2. If $p_{i,j} \leq p_{k,j}$ & $d_{i,j} \leq d_{k,j}$, b_i can provide better call quality than b_k . We use $b_i : (p_{i,j}, d_{i,j})$ as a new benchmark to continue base station selection.
3. If $p_{i,j} \geq p_{k,j}$ & $d_{i,j} \leq d_{k,j}$ or $p_{i,j} \leq p_{k,j}$ & $d_{i,j} \geq d_{k,j}$, it is hard to judge directly which one is better. The MOS values will be computed and used for base station selection.

Let B_j be the set of base stations which satisfy the bandwidth requirement for the mobile terminal m_j . The detailed algorithm for base station selection is given in Algorithm 1.

Let $|B_j|$ be the number of base stations in B_j . The time complexity of Algorithm 1 is $|B_j|$. The proposed MOS-based base station selection scheme has several advantages in comparison with the existing solutions. Firstly, the algorithm guarantees that the selected base station must meet the requirement on bandwidth, thus avoiding frequent handover failures as in simple RSS-based solutions. Secondly, the proposed scheme is energy efficient since there is no need

```

input :  $B_j = \{b_i : (p_{i,j}, d_{i,j})\}, b_k, m_j$ 
output: optimal  $b^*$ 

 $p = p_{k,j}; d = d_{k,j}; M = M_{k,j}; b^* = b_k;$ 
for each  $b_i : (p_{i,j}, d_{i,j})$  in  $B_j$  do
    if  $(p_{i,j} \geq p) \& (d_{i,j} \geq d)$  then
        | continue;
    else if  $(p_{i,j} \leq p) \& (d_{i,j} \leq d)$  then
        |  $p = p_{i,j}; d = d_{i,j}; M = M_{i,j}; b^* = b_i;$ 
    else if  $(p_{i,j} \geq p \& d_{i,j} \leq d) \text{ or } (p_{i,j} \leq p \& d_{i,j} \geq d)$  then
        | Calculate  $M_{i,j}$  ;
        if  $(M_{i,j} \geq M)$  then
            |  $p = p_{i,j}; d = d_{i,j}; M = M_{i,j}; b^* = b_i;$ 
        end
    end
end

```

Algorithm 1: Optimal Base Station Selection

for continuous scanning. In RSS-based solutions, the mobile terminal has to continuously scan the current available networks, which consumes quite a lot battery power. In our scheme, the RNC will be responsible for collecting QoS parameters and making handover decisions. Thirdly, our scheme can avoid unnecessary handover. In situations where a user is having low RSS but the quality of voice call from a user's point of view is acceptable, the RSS-based handover protocol takes decision to do handover even if there is no need to do so.

The proposed algorithm is very efficient when the mobile user intend to do the handover and there are multiple base stations available for handover. We only process if the bandwidth criteria is met, which results in less computation and reduces the complexity of the system. This is very efficient in the scenario where RNC has to handle multiple handovers at a time. It is also a QoS aware scheme which guarantees the quality to the user unlike other schemes which are based on only RSS, load, power, etc.

There are problems with handover protocols based of RSS, the mobile node has to continuously scan for the current and available networks. Scanning uses battery power, and is not energy efficient if mobile node has to scan continuously. The

handover protocol which we designed is energy efficient. The mobile node does not have to scan for the available network continuously. The protocol which is located at the network receives the QoS parameter and depending upon these parameter decision of scanning the available base station is taken. Only if these parameters are below threshold we decide to scan for the other available network. This also avoids unnecessary handover as there might be situation where user is having low RSS but quality of voice call from the user point of view is good in which case the RSS based handover protocol takes decision to do handover even if there is no need to do handover as quality of the signal is acceptable.

4.5 Handover Protocol Design

This section presents our MOS-based handover protocol of which the design is based on the IEEE 802.21 framework (IEEE, 2008). We will first give a brief overview of the IEEE 802.21 standard, and then describe the details of our handover protocol.

4.5.1 IEEE 802.21 Standard

The IEEE 802.21 standard defines a media-independent handover (MIH) framework that can significantly improve seamless handover between heterogeneous network technologies. IEEE 802.21 facilitates the handover between different radio access technologies without call interruption, providing seamless connectivity for the mobile terminal, and improving the quality of service. MIH framework is based on a protocol stack implemented in all the devices involved in the handover, and provides a common interface for the link layer functions which are independent of radio access technologies. It consists of a MIH client which sits at user equipment end. MIH server resides in the core network. Handover decision for all the users in that zone is based on the information provided by MIH. In the IEEE 802.21 standard, Media Independent Handover Functions (MIHFs) are defined to provide a generic link layer.

The MIH framework provides a group of MIH functionalities that facilitate both mobile-initiated and network-initiated handovers. MIH provides a framework which exchanges the events, commands and information about QoS parameters,

current link layer conditions and traffic load with different radio access technologies, which are used as input for taking decision for handover. The major components include:

- MIH function (MIHF), which is a logical entity that provides abstract services to the higher layers through a media independent interface and obtains information from the lower layers through media specific interfaces. It provides three types of services: (1) Media-Independent Event Service (MIES) for detecting and reporting changes in link layer properties; (2) Media-Independent Command Services (MICS) for local or remote MIH users to control link state; and (3) Media-Independent Information Service (MIIS) for providing information about neighboring networks.
- Service Access Points (SAPs), which define both media-independent and media-specific interfaces. It includes: (1) MIH_SAP for high layers to control and monitor different links; (2) MIH_LINK_SAP for MIHF to control media-specific links; (3) MIH_NET_SAP to support the exchange of MIH information and messages with the remote MIHF.

For the details of the framework, refer to (IEEE, 2008)(Taniuchi, Ohba, Fajardo, Das, Tauil, Cheng, Dutta, Baker, Yajnik and Famolari, 2009).

4.5.2 Parameter Acquisition

To perform the MOS-based handover, our protocol needs the following information: the set of candidate base stations, the available bandwidth capacity at each candidate base station, the delay and packet loss rate for each wireless link between the mobile terminal and candidate base station, and the MOS of the current connection. All these information can be obtained in the following way.

Candidate base stations: The neighboring base stations information can be collected using the Media-Independent Information Service (MIIS) in IEEE 802.21. The intelligent MIH connection monitoring manager sits between the application and the device radio modem to monitor the wireless access, network status and availability. Link manager is responsible for managing local link. It controls the local link by responding to MIH commands.

Bandwidth: Each base station keeps track of its available bandwidth capacity. As all base stations are wired to the RNC, the available bandwidth capacity of

each base station can be reported to the RNC at a regular time interval, or can be retrieved by the RNC dynamically.

Delay and packet loss: When a VoIP call is made from any device (mobile, laptop and iphone etc.), it travels through the mobile terminal, NodeB (base station), RNC, SGSN, Gateway GPRS Support Node (GGSN) to the Internet. The only air interface is between the mobile terminal and the base station, which commonly has a significant effect on the call quality. In a heterogeneous network, each mobile terminal is equipped with multiple radio receivers. The quality of the links from the mobile terminal to different networks can be monitored using the MIES and the SAPs functions. The link manager detects the link quality of the current call. Connection monitoring manager detects the link quality of the available base station. To calculate the value of delay, packet loss and MOS of the candidate base station, there are two different approaches: 1) Calculate the packet loss, delay and MOS by sending test packets from the mobile terminal to other available base stations. 2) Establish multiple tunnels between the mobile terminal and the base stations at a time (Zahid Ghadialy, 2007). With the multiple tunnels approach, a mobile terminal has to establish a tunneling with the available base stations. We use the first approach because the multiple tunnels approach causes an extra overhead on the mobile terminal.

MOS: The MOS value of the current ongoing call is used to initiate the handover process and to compare with other potential connections. This information can be calculated as there is an ongoing connection between the mobile terminal and the serving base station. For the MOS values of other potential connections, they can be estimated based on the delay and packet loss rate using Equation (4.6).

4.5.3 MOS-based Handover Protocol

In our protocol, the handover is triggered by the mobile terminal, whereas the decision on whether the handover will be finally performed and how the handover is performed is made at the RNC. The mobile terminal monitors the call quality and the current link state. If the mobile terminal detects the MOS value of the current ongoing call is below a predefined threshold, it sends a request to the RNC for handover. Once the RNC receives the request message, it will send a query to the candidate base stations. If there is another base station that can provide better service to the mobile terminal, handover will be immediately executed;

otherwise the handover request is rejected. Figure 4.5 shows the flowchart of the protocol we designed, which consists of three steps: handover request, base station selection, and handover execution.

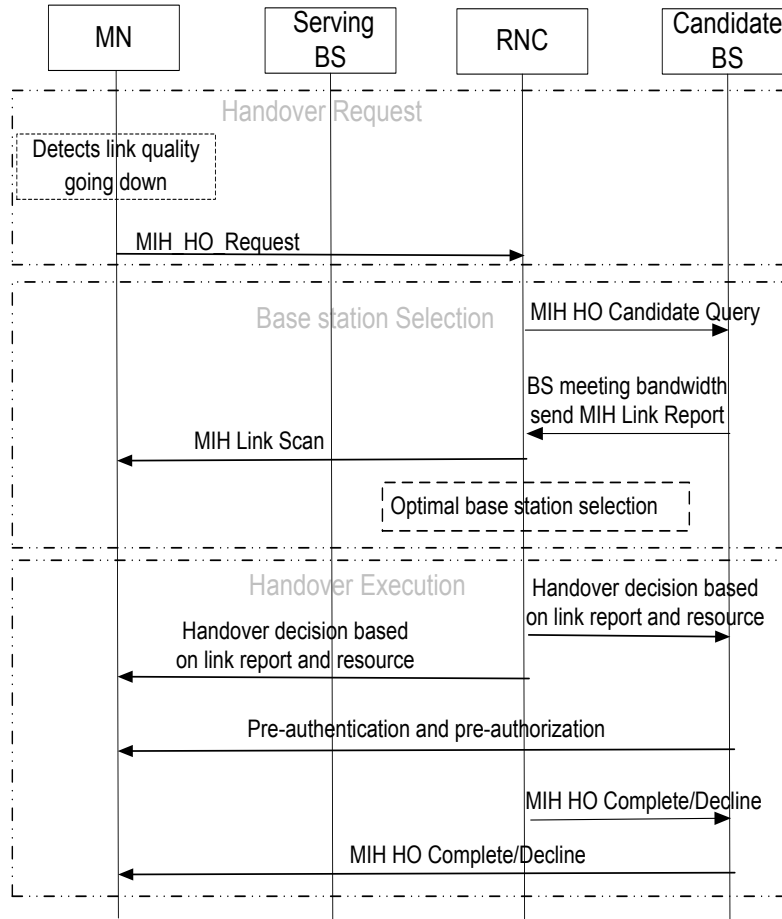


Figure 4.5: Handover Design Protocol

Handover Request: When a mobile terminal m_j served by the base station b_i detects that the MOS value of the current ongoing VoIP call is below the threshold, it sends an MIH_HO_Request message to the RNC, and this message contains the following information: (1) the current MOS value $M_{i,j}$ (2) the current link status $\{p_{i,j}, d_{i,j}\}$; and (3) the bandwidth requirement c_j .

Base Station Selection: Once the RNC receives an MIH_HO_Request message from mobile terminal m_j , it broadcasts an MIH_HO_Candidate_Query message encapsulating the required bandwidth c_j to all the base stations. Only the base stations which have available bandwidth no smaller than the requested bandwidth c_j , will join the link measurement process. After link quality measurement,

each base station b_k which has available bandwidth no smaller than c_j sends an MIH_Link_Report message to report $(p_{k,j}, d_{k,j})$ to the RNC, and b_k will reserve the bandwidth c_j for mobile terminal m_j . The RNC also sends an MIH_Link_Scan message to the mobile terminal m_j to initiate the process for measuring the quality of the wireless links from the mobile m_j to the other base stations. When the RNC collects all the link reports from the suitable base stations, Algorithm 2 is executed to compute the optimal base station for handover.

Handover Execution: If there is no other base station that can provide better call quality than the current serving base station, the RNC sends an MIH_HO_Decline message to the mobile terminal to terminate the handover process, and sends another message to the other base stations to release the bandwidth resource reserved for the mobile terminal m_j ; otherwise the detailed steps are performed for handover from the current serving base station to the new base station. When the handover is completed at the higher layers, a MIH_HO_Complete message to the MIH.

The handover process involves measurement of the parameters, calculation of the constraints, starting the handover process, comparing the threshold parameters with the obtained and based on these take the decision of handover. We propose vertical handover protocol for a moving user between UMTS and WiMAX network or vice a verse considering a VoIP application. Proposed VHA is network based which is located at RNC. We use IEEE 802.21 standard which is MIH framework to obtain the input to VHA. MIH is used to facilitate the handover by proving a common language between heterogeneous networks. IEEE 802.21 can initiate the handover mechanism between different radio access networks by providing the capability for obtaining the necessary information for handovers.

Our protocol needs information of the input parameters which are bandwidth, MOS, delay, packet loss, and the candidate base stations available for the handover. Figure 4.5 shows the flowchart of the protocol we designed. As seen from the figure, we monitor the MOS for the mobile equipment connected to the current base station. MOS value of the current call is calculated at the mobile node using Equation (4.6). If the mobile equipment detects MOS value of the current ongoing call is below threshold, it sends the MIH_Scan signal to MIH which scans for the candidate base station. Then MIH_LINK_SAP, a media specific Service Access Point (SAP) provides an interface for the MIHF to control and monitor

media specific links. `MIH_Get_Status` gets the status of link. `MIH_Link_Report` provides link reports. The TE then sends `MIH_Handover_Initiate` signal to the RNC which initiates the handover process. It also sends the information of the MOS value and candidate base stations available for the handover. RNC sends acknowledgment message to mobile equipment.

Once the initial parameters for the handover are obtained, constraint on the bandwidth is checked. The RNC broadcasts the information of the required bandwidth to the candidate base station. The required bandwidth for the application of the current ongoing call is known to the base station which is serving the user. This is calculated before the serving base station accepts the call. Only those base stations satisfying the bandwidth criteria acknowledge the RNC. Base stations which do not have sufficient bandwidth are not considered for the further calculation. Accepting the call without having sufficient capacity might result in degradation of call quality and user might again have to do handover resulting in the unnecessary handovers.

RNC then sends the information of the available base stations to the mobile equipment for the calculation of the packet loss and delay. These parameters are calculated from the information of the packets received. When a VoIP call is made from any device (mobile, laptop, iphone etc) it travels through mobile equipment, NodeB (BS), RNC, SGSN, Gateway GPRS Support Node (GGSN) and internet. The only air interface is between mobile equipment and Node B, rest all the connection is wired. For our study we consider packet loss and delay from mobile equipment to BS which is the only air interface where quality of call affects significantly. To calculate the packet loss and delay of the candidate BS we send beacons from the mobile equipment to the BS. The information of the packet loss and delay is then send to RNC, where intial comparison is done according to the observation. If the observation fails we calculate the MOS value. This is done for all the base base stations which meets the bandwidth criteria and depending on these condition decision of handover is taken. If all these conditions meet the criteria handover is done otherwise mobile stays connected to the same network avoiding unnecessary handover.

To evaluate the proposed scheme, we implemented our MOS-based handover protocol in MATLAB. We simulated it in an integrated environment with WiMAX and UMTS networks, and compared its performance with the RSS-based han-

dover (Sarddar, Maity, Raha, Jana, Biswas and Naskar, 2010). In this study, we measure MOS, Handover Dropping Probability (HDP) and Call Dropping Probability (CDP).

4.6 Simulation Setup

To evaluate the proposed model we designed and implement the MOS based handover protocol in Matlab2011. We simulated an integrated environment of WiMAX and UMTS networks. We assume the mobile user is using the VoIP application. In our simulation, we deploy 10 UMTS base stations and 4 WiMAX base stations in a $10000m*10000m$ area. The mobile terminals are uniformly placed in the UMTS or WiMAX cells. Each of the UMTS or WiMAX cells has a base station. All the base stations are connected to the RNC where the handover algorithm is located. The diameter of a UMTS cell is configured to 2 km, and the diameter of a WiMAX cell is configured to 3 km. When a mobile terminal makes a VoIP call, the voice packets are carried from the mobile terminal to the RNC through Node-B. Even though there are different codecs, such as G.711, G.721 and G.722 etc. We use G.711 since it has the least compression delay (Cole and Rosenbluth, 2001). Each simulation is run for 10 minutes.

In our simulation, we use a 2D random walk model to simulate the movement of the mobile terminals. Because some mobile terminals are believed to move in an unexpected way, random walk mobility model is proposed to mimic their movement behavior (Bai and Helmy, 2004). The random walk model is a stateless mobility process, where the information about the previous status is not used for the future decision. That is, the current parameter information is independent with its previous parameter information.

The movement of each mobile terminal is controlled by two parameters: the moving direction θ and the step size L . Each time the mobile terminal steps for the distance $L = 0.1$ meters, a new direction is randomly chosen (i.e. four possibilities: 1) forward, 2) backward, c) left and d) right) then steps for another distance L randomly. In this study UMTS and WiMAX networks are used and the application used is VoIP. The simulations were started at the same time. The behavior of single call and 100 calls are studied. The performance changes with the amount of the traffic. The problem formulation shows the base station

selection problem is formulated with capacity as constraint.

In our simulation, the capacity is taken as a constraint for the base station. We do not take it as a measuring parameter. We take this decision because the simulation results does not affect by the capacity, since MOS does not get affected by the capacity. We consider capacity because the new base station to which the call is going to be handed over needs enough capacity to accept the call else it will be rejected. Capacity of every base station is decided by the operator depending on factors such as the number of people living in the residential, business area or depending on the geographical positions. However, its very hard to predict number of users near all base stations due to users random mobility. Keeping this in mind, in our set up we have assumed random initial capacity for every base station and the capacity varies dynamically with the number of users attached to respective base stations. In realistic scenario the user might run VoIP keeping email or SMS in background. The background traffic such as email are not real-time traffic and the preference will always be given to real-time applications. The delay-tolerant applications are served by network based on as and when network has capacity. These applications will impact the network performance, however the impact is very minimum. The non-real time applications are of less priority compared to real time applications like VoIP. If the network capacity has reached its maximum these non-real time applications will be queued up and served when channels are available. The reason for selecting the VoIP application in our study is that VoIP is a combination of voice and IP and is becoming very popular because of its low cost.

We compare our MOS-based handover protocol with RSS-based handover protocol proposed in (Grønsund, Grøndalen, Breivik, and Engelstad, 2007). The RSS is calculated using the following function

$$RSS = -62.5 - 26.5 * \log_{10}(d) \quad (4.11)$$

where d is the distance in kms between a mobile terminal and a base station. If mobile terminal detects the RSS value below the threshold, the mobile terminal scans for the available networks and hands the call over to the base station providing higher RSS. If the mobile terminal fails to find a better base station, the handover request is rejected, and the mobile terminal continues the call by connecting to the same base station.

4.7 Simulation Result

In our simulation, we use the following four metrics to evaluate the performance of WiMAX and UMTS in terms of end-to-end QoS for VoIP. We measure critical parameters such as MOS, Call Dropping Probability and handover dropping probability. In this section we present the experimental results of MOS based handover scheme. We compare the results of MOS based handover scheme with RSS based scheme and explain the significance of the MOS based handover scheme. Both the schemes are simulated with resource capacity constraints. Quantitative analysis of MOS, HDP and CDP are presented.

4.7.1 Mean Opinion Score (MOS)

In this set of simulations, we use only one mobile terminal and monitor the MOS during its movement. The MOS threshold for MOS-based handover scheme is set to 3.0. Figure 4.6 shows the MOS values for the proposed MOS-based handover scheme. The handovers occurred are marked by the numbers on the graph. Initially, the mobile terminal has a MOS value of 3.6. Until the end of the first minute, the mobile terminal maintains the connection with the current serving base station. The mobile terminal detects the decline in quality after the first minute and performs a search operation looking for the base station with the highest MOS among the available base stations for handover. Since the mobile terminal does not have a base station with a stable and better MOS value, it continues the service with the current base station until it finds a suitable base station at approximately the second minute. Once the mobile terminal is able to get a base station with a stable MOS value higher than the threshold, it performs a handover which is numbered 2. From the graph, we can see that the MOS value of the call is maintained and the mobile terminal tends to select the base station which provides better MOS value every time it decides to make a handover due to degradation of the call quality, and thus guarantees the call quality.

Figure 4.7 displays the MOS value for the RSS-based handover scheme, where the RSS threshold for handover is configured to -68dBm (Grønsund, Grøndalen, Breivik, and Engelstad, 2007). RSS-based and MOS-based handover schemes are simulated in the same scenario (i.e., same movement, same environment). The mobile terminal starts with an initial value for MOS of 3.6 and continues the

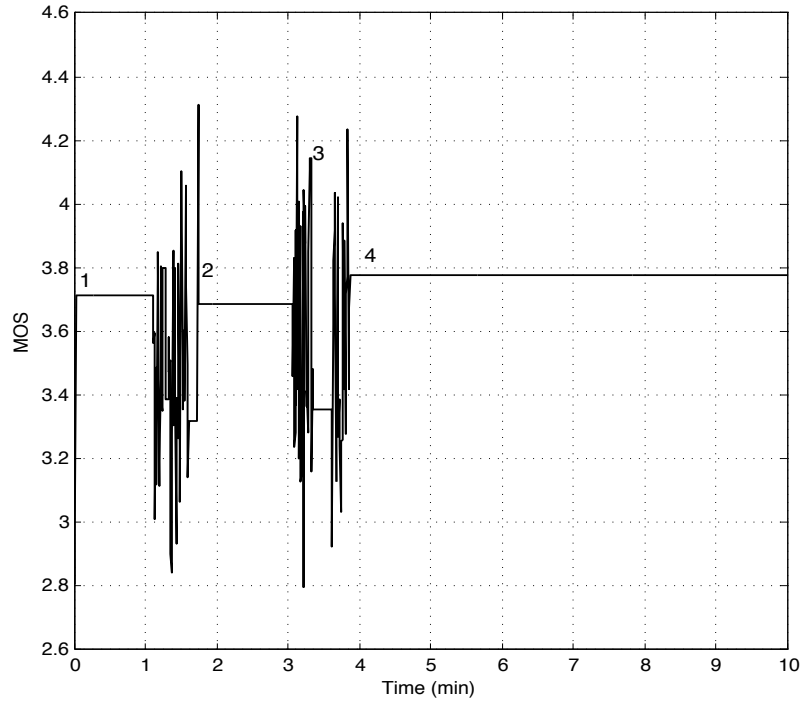


Figure 4.6: MOS-Based Scheme

connection with the current base station until it realizes a drop in the RSS value below the threshold. The mobile terminal performs a scan for the target base stations with better RSS. Since the RSS-based handover scheme do not consider the MOS value of the target base station, the probability of choosing a base station with a larger RSS but a smaller MOS value is higher. For example, it can be seen from the figure that the first handover of the mobile terminal is performed at approximately the fourth minute by selecting a base station with the MOS value of 2.5. As the MOS value is lower, the user experiences a poor call quality. After the fifth minute the mobile terminal again experiences a poor RSS and performs a handover. This time the MOS value of the base station is slightly better than the previous one but not the best one to service the call.

It can be seen from Figure 4.6 and Figure 4.7 that MOS-based handover scheme provides better MOS value than the RSS-based scheme. Since the MOS value is a QoS parameter which is calculated from the delay and packet loss, this indicates that the new MOS-based handover scheme is better than the RSS-based scheme in terms of maintaining quality of the call. The drawback of our MOS-based scheme is, the processing time at the RNC is more than the RSS-based scheme, as it involves the collection of delay and packet loss parameters, and then, calculation

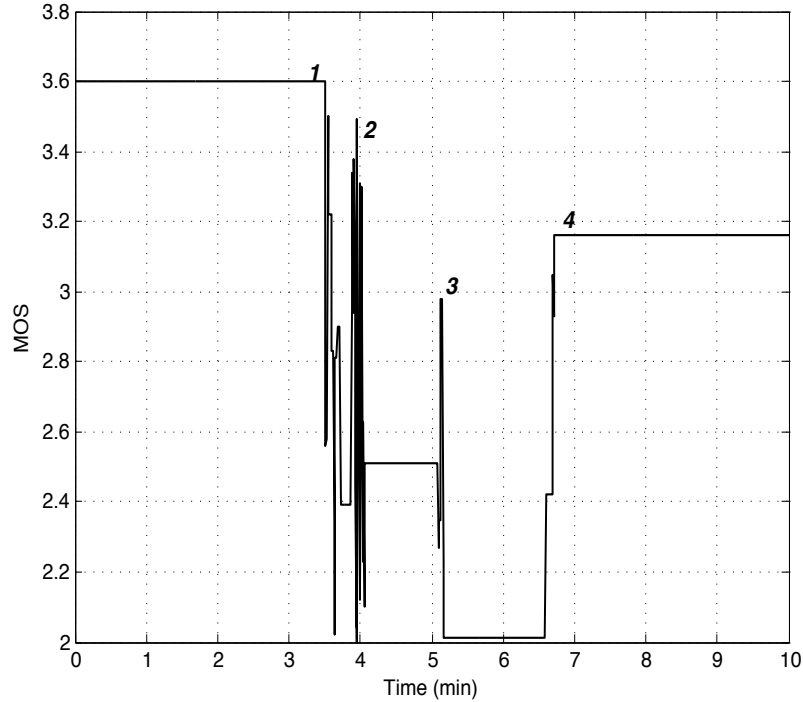


Figure 4.7: RSS-Based Scheme

of the MOS for the candidate base station once the handover request is made. The above simulation was ran for 10 times and it was observed that the average MOS in MOS-based handover scheme was 3.9. The average MOS for RSS-based handover was 3.4.

4.7.2 Handover Dropping Probability (HDP)

When a mobile terminal requests for a handover, the handover process is dropped if a handover request is not processed. The corresponding probability is called as handover dropping probability. The call still continues with the current attached base station. The dropping probability is given by $p_d = x/y$, where x is the number of unsuccessful handover and y is the number of handover requests (Shuaibu, Syed-Yusof and Fisal, 2007). HDP is a QoS metric and can be used as a performance indicator of a system. When a mobile terminal moves from one cell to another cell or from one network to another network, the call has to be transferred without dropping or degrading the quality. For a mobile terminal to maintain seamless connectivity, HDP plays a very important role.

For both MOS and RSS based schemes we calculated HDP for 100 users. Figure

4.8 shows handover dropping probability in MOS-based and RSS-based handover schemes. The HDP of RSS-based handover is higher than the MOS-based handover schemes. The handover process fails for various reasons, such as no sufficient bandwidth, no enough wireless resources to provide strong signal strength to the mobile terminal receivers, more packet loss and delay in turn affecting MOS or could not support the application well. The handover dropping probability depends on the type of handover schemes. In the case of RSS-based handover, Line of Sight (LOS) plays an important role in the success of the handover process. If a mobile terminal is not in the LOS it affects the RSS. When a user senses low signal strength and invokes a handover request, the possibility of high rate of rejection is true which in turn increases handover dropping probability. During the handover process, the packets are buffered at the mobile terminal and at the base station. Due to the limited buffer size or buffer overflow, the handover request is put in the queue and serviced according to the First-In-First-Out procedure. The handover is dropped if the waiting time expires. If the mobile terminal is moving in the heterogeneous network, for RSS-based handover the user equipment has to have multiple antennas to detect the signal strength from multiple networks. But in case of our handover scheme, it is not necessary for a mobile terminal to have multiple antennas since we use the MIH protocol to detect the link layer state. The above simulation was ran for 10 times and it was observed that the average HDP in MOS-based handover scheme was 0.23. The average MOS for RSS-based handover was 0.27.

4.7.3 Call Dropping Probability (CDP)

When a handover request cannot be processed and the call cannot be serviced either by the current base station or the candidate base station, the call is dropped and the probability is called call dropping probability. The call dropping probability is given by $p_c = x/y$, where x is the number of unsuccessful calls and y is the total number of call requests.

CDP is not to be confused with HDP. These are two different parameters but are inter-related. When a mobile terminal requests for handover, the request can be granted or denied. When handover request is denied, the handover process is dropped. The corresponding probability is called handover dropping probability, the call still continues on the current base station. But in case of CDP, the call is

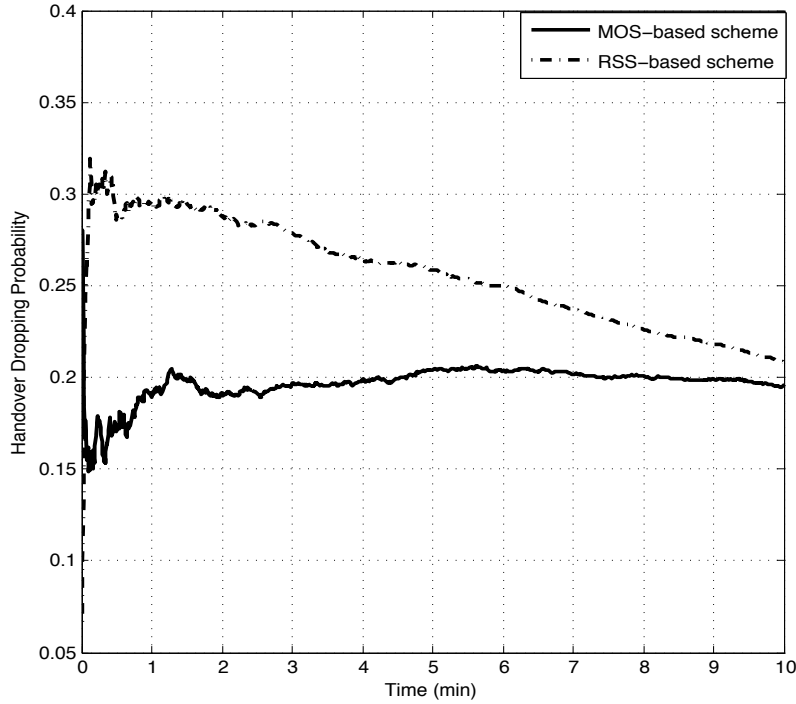


Figure 4.8: Handover Dropping Probability

dropped if it cannot be serviced by any of the base stations. CDP is a subset of HDP but not vice versa. The call dropping rate is used as an indicator to identify network congestion or to realize the base station has a higher packet loss and delay. Call dropping is more serious than handover dropping. HDP and CDP can be caused by poor signal strength, scarce wireless resources, wireless transmission delay and wireless channel access delay. It increases as the congestion in the network increases or due to poor reception caused by the fading signal. The more the number of handovers per base station, the more will be the buffering of packets at the base station, which in turn increases the HDP and CDP. The above simulation was ran for 10 times and it was observed that the average CDP in MOS-based handover scheme was 0.01. The average MOS for RSS-based handover was 0.025.

For both MOS and RSS based scheme we calculated CDP for 100 users. Figure 4.9 shows call dropping probability in MOS-based and RSS-based handover schemes. RSS-based scheme has higher call dropping probability than the proposed MOS-based scheme. It can be seen from the Figure 4.8 that in RSS-based scheme more handover requests are dropped. Since HDP is higher and if mobile terminal cannot be serviced by the current base station, the call is dropped as well. As

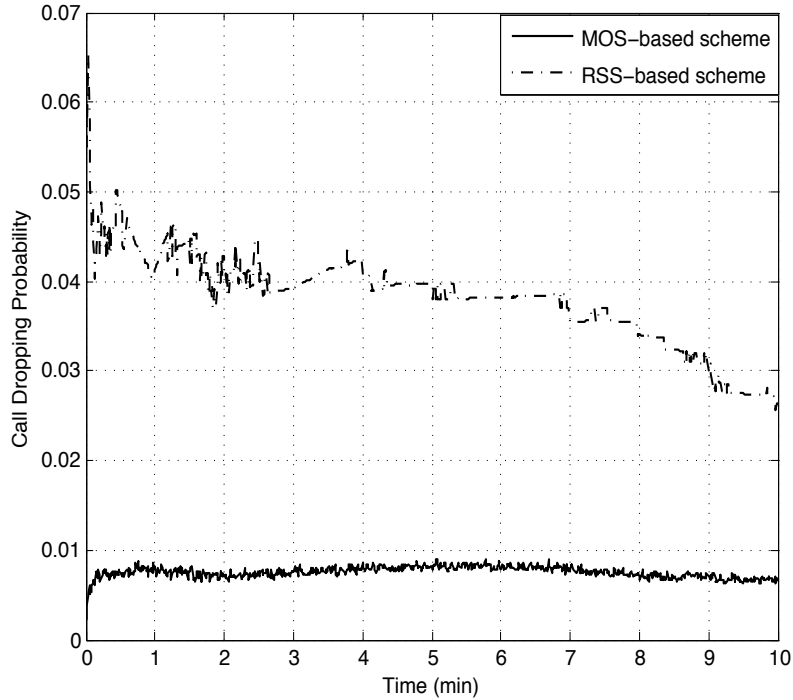


Figure 4.9: Call Dropping Probability

seen from Figure 4.9, MOS-based handover has a lower call dropping probability than the RSS-based scheme. When the current base station cannot service the call, the mobile terminal requests for a handover. When the mobile terminal makes a request for handover, the base station has to minimize the number of dropping calls. The call is dropped if the mobile terminal cannot be serviced by a candidate base station as well as the current base station.

In the case of RSS-based handover, the CDP is higher because of higher HDP. In RSS-based handover when a handover is initiated, and if the candidate base station is not able to service the call due to LoS or capacity the handover gets dropped and the call continues with the current base station. The current base station will not be able to service the call as the user is moving away from the current base station. As the user moves away the signal strength reduces and the call is dropped.

4.8 DISCUSSION

Handover is necessary when a connection needs to be transferred between cells or networks for seamless connectivity and good QoS. In this work, we proposed

a novel handover scheme compliant with the IEEE 802.21 standard, that enables a wireless access network to transfer the call between cells or networks, taking care of the quality of the call and load among all the available base stations. We formulated the base station selection problem as an optimization problem with the objective to maximize the call quality, and presented a scheme to forward data packets to the most appropriate base station in order to maintain good quality of the call. We conducted extensive simulation using a scenario of urban network environment with VoIP call in WiMAX and UMTS integrated networks and analyzed critical QoS parameters like MOS, CDP and HDP. We compared our proposed scheme with the RSS-based handover scheme. Results show that our proposed scheme provides higher MOS values, thus improving the perceived quality of the call and reduces the HDP and CDP. It is a QoS aware scheme which guarantees the call quality to a user. The proposed scheme is also energy efficient as it does not require to scan the network frequently. The energy preserving issue is going to be addressed in the next chapter.

Chapter 5

Energy Efficient Handover for NGWN

5.1 Introduction

Wireless communication is spreading at a tremendous speed and becoming a vital part of the modern communication. Recently, most of the devices have turned into wireless, and supported multimedia applications such as video streaming and VoIP which consume a considerable amount of energy. However, most state-of-the-art wireless devices like smartphones and laptops are powered by battery. Battery power has become one of the most valuable and limited resources for this category of devices. The size of the NGWN devices are reducing and they are expected to perform more efficiently, e.g. size of the battery is being reduced, and it is expected to last longer. Since the mobile terminals are dependent on battery power, it is extremely critical to minimize their energy consumption. Battery life can be prolonged by reducing the data rate, which in turn affects the quality of the call. There is a trade-off between the communication quality and energy consumption. Energy consumption has been a major issue on battery powered handsets like cellphones due to the expectation to make the devices operate as long as possible. Thus, it is essential to develop efficient solutions to minimize the energy consumption while still maintaining satisfactory communication quality.

Handover is a process of transferring a call from current base station to another to provide seamless wireless connection. To perform handover, a mobile terminal needs to first scan all the available channels. Moreover, handover may be dropped

due to some unsatisfied constraints like capacity at the base station. To continue the handover process, channel scanning must be re-done, thus consuming quite a lot of energy at the mobile terminals. Another major part of energy is consumed by wireless data communication. During the handover process, the mobile terminal may unfortunately choose a base station which requires the mobile terminal to transmit at a high power level to guarantee satisfactory communication quality, thus consuming a significant amount of energy. It can be seen that handover plays a prominent role in saving energy at the mobile terminal side, and an energy-efficient handover solution should try to minimize energy consumption for both channel scanning and data communication. The traditional handover schemes (Kunarak and Suleesathira, 2010) commonly use the RSS as the metric for base station selection. However, RSS alone is not sufficient to make handover decision as it does not take into consideration factors like quality and energy.

There are energy saving techniques proposed at various network layers. The authors in (Jones, Sivalingam, Agrawal and Chen, 2001), (Karl, Holger and Willig, Andreas, 2007) gave a survey on energy efficient and low-power design at all the layers of wireless network protocol stack.

- Physical Layer: Physical layer handle functions such as modulation, demodulation, channel coding and encoding. It mainly deals with the hardware side of the mobile terminal. In order to reduce the energy consumption at the physical layer many low-power design circuits have been proposed. There are new modulation techniques such OFDMA which increase the capacity and reduce the energy consumption.
- Data link Layer: It preforms functions, such as allocating the bandwidth, encryption, decryption, error detection and error correction. The data link layer provides a mean to transfer data between the different network components. To reduce the energy consumption at data link layer, efficient error detection and correction algorithms are proposed. There are techniques proposed at the MAC layer to minimize re-transmissions.
- Network Layer: This layer is responsible for routing packets, sending packets from the data link layer to transport layer and mobility management. In order to reduce the energy consumption at this layer, various energy efficient handover schemes are proposed. There are various routing techniques such as shortest path proposed to reduce the energy consumption.

- Transport Layer: The transport layer is responsible for delivering the data efficiently and reliably to the required application. It is also responsible for congestion control. Various energy efficient congestion control mechanisms are available at the transport layer, in order to avoid the loss of packets and recovery.
- OS/Middleware Layer: This layer manages QoS and power in the network. It also manages access to the physical resources such as CPU in the network.

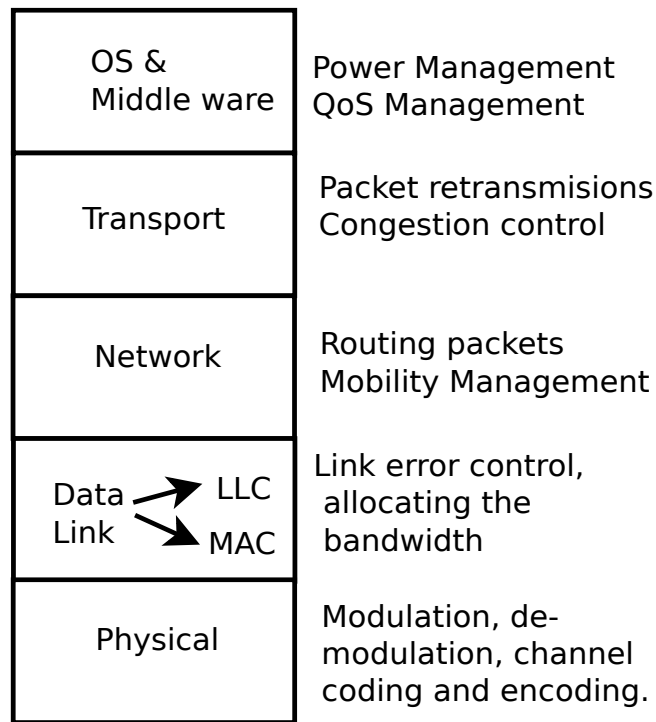


Figure 5.1: Energy saving at different layers

In this chapter, we address the issue of designing an energy-efficient handover scheme to minimize the energy consumption at the mobile terminal subject to the constraint on communication quality. We take into account the energy consumption for both data communication and channel scanning. To save energy for wireless data communication, we compute the minimum transmission power needed for each channel to provide the desired QoS. For channel scanning, we use the handover dropping probability to estimate the energy consumed by scanning. By formulating the handover as an optimization problem, we propose both a centralized solution and a heuristic solution for the base station selection. An energy-efficient handover protocol is then designed based on the IEEE 802.21

standard, which is also called MIH.

The rest of the paper is organized as follows. Section 5.2 briefly discusses the related work. Section 5.3 presents the system model. Section 5.4 gives the problem formulation. Section 5.5 describes the proposed solution. Section 5.6 presents the handover protocol design. Section 5.7 discusses the simulation setup. Section 5.8 discusses the performance metrics. Section 5.9 discusses the numerical results. Finally, in Section 5.10 we conclude with a discussion.

5.2 Related Work

Some related work includes schemes on minimizing the energy consumption with delay constraint. The authors in (Prabhakar, Uysal Biyikoglu and El Gamal, 2001) investigated the packet transmission scheduling that reduces energy subject to a deadline or a delay. An algorithm called lazy online algorithm which varies transmission time according to packet accumulation was designed. The work in (Uysal-Biyikoglu, Prabhakar and El Gamal, 2002) considered packet transmission scheduling that reduces energy subject to a deadline or a delay constraint by altering the packet transmission time and power levels. An algorithm to compute the optimal schedule for transmitting the packets within the given amount of time was proposed. However, this strategy might not be practical, as transmitting a packet over a longer period of time might incur additional delay, or cause packet loss. As a result, this might involve sending the packet again causing more energy loss. In particular for real time applications which are time sensitive, adding an extra delay will significantly affect the QoS.

In (Balasubramanian, Balasubramanian and Venkataramani, 2009), the authors conducted a detailed evaluation research and found significant energy overhead in Wi-Fi, 3G and GSM networks. For each technology, the authors developed a measurement driven model of the energy consumption. Some existing energy-aware handover schemes focus on adding an extra knob which can switch between different codec or adjust the modulation schemes at the physical layer to conserve energy. Solutions based on switching the active network interface have also been proposed (Nam, Choi, Seok and Choi, 2004). However, switching the network interfaces incurs additional delay and requires additional energy. According to the authors in (Feeney and Nilsson, 2001), most of the energy-conserving link-layer

protocols concentrate on the centralized base station approach. These protocols rely on a resource-rich base station and allow limited mobile terminals to spend time in low-power consumption sleep state. However, these strategies are not feasible in the environment where a mobile terminal is moving and has no fixed base station.

The authors in (Choi and Choi, 2007a) presented an energy aware scanning scheme in an integrated IEEE 802.16e/802.11 network. It was proposed to do modifications in 802.16e network such that each base station broadcasts the density of WLAN access point residing within its cell coverage. This strategy requires periodic broadcasting of information about the density, incurring additional overhead of resending the packets which consumes more energy. In (Yang, 2007) an algorithm was proposed to dynamically change the power by adjusting the timer threshold and discontinuous reception cycle values to improve the performance. The work in (Verdu, 2002) showed that the bandwidth-power trade-off for a general class of channels in the wide-band regime is characterized by low, but nonzero, spectral efficiency and energy per bit close to the minimum value required for reliable communication. In (Wang, Wang and Nilsson, 2006), the authors investigated the relationship between the energy consumption and transmission rates of the mobile terminal, and proposed a scheme to reduce energy consumption of each terminal. The work in (Yang, Wang, Tseng and Lin, 2009) takes into account the geographic mobility and proposed a handover scheme considering the past handover patterns of mobile terminals. Practically it is not feasible as a user's movement is random. Therefore, making a handover from past handover pattern might lead to choose a base station which does not provide adequate quality to the user.

In (Claussen, Ho and Pivit, 2008) the authors proposed a method of using joint macro and pico cell coverage to increase the energy efficiency of cellular networks. The authors showed that an increase in the deployed number of femto cells will result in the reduction of total energy consumption. However, the authors do not consider the deployment cost and complexity involved in the femto cells.

5.3 System Model

In this section, the system model for minimizing the energy consumption of the mobile terminal during handover and energy consumption for a communication link by using quality constraints is presented. When the mobile terminal requires a handover of call from one base station to another base station, it consumes a significant amount of energy, which is spent on scanning the new links, communication link and overhead caused in mapping the parameters from one network to the another network. In this study the infrastructure network is considered, and the VoIP application is used. The problem is divided into two parts: 1) Energy during the link communication, where the optimal base station link is selected for the mobile terminal, which requires minimum transmission power. 2) Energy during scanning, where MIH is used to collect the link information. These two objective functions are combined to minimize the total energy consumption during the handover process in the integrated network.

An example of overlapping WiMAX and UMTS networks deployed in an urban area as shown in Figure 5.2 is considered. However, our proposed scheme can be applied to any homogeneous or heterogeneous networks like UMTS, WiMAX, Wi-Fi or LTE. In our model, handover is initiated by the mobile terminal and is executed at the handover controller. The loosely coupled architecture is used (Xu, Zhang, and Zhou, 2007a), in which a handover controller maintains the network information such as the quality of UMTS network and WiMAX network. Handovers between these two integrated networks are managed and optimized by the handover controller which is connected to the RNC and ASN gateway. Let $B = \{b_1, b_2, \dots, b_n\}$ denote the set of UMTS and WiMAX base stations deployed and $U = \{u_1, u_2, \dots, u_n\}$ represent the set of mobile terminals in the urban area.

5.3.1 Energy on Data Communication

With growing applications on the mobile terminal, energy consumption has increased. These applications require high transmission power due to high data rate and bandwidth. The required energy depends on many factors, such as distance between the mobile terminal and base station, SINR environmental factors and interferences. Energy in the wireless communication is a function of the transmission power used to operate the successful link connection. Cost of a link is the

where t is the communication time (second). The unit of $P_{i,k}^T$ and $P_{i,k}^C$ is joules/sec or dBm.

There has been substantial amount of energy consumed at the circuit level and researchers are working on low power circuit design. In this work we do not consider the energy communication circuit used $E(P_{i,k}^c)$ during the active communication as it generally remains constant. Unlike wired networks, wireless links are lossy and the link quality might change over time. To achieve high transmission reliability, sometimes the transmission power has to be increased. The energy required for reliable transmission over link (i, k) , denoted by $E_{i,k}^{RT}$, is given by (Banerjee and Misra, 2002) (Banerjee, Suman, Misra and Archan, 2004),

$$E_{i,k}^{RT} = \frac{P_{i,k}^T}{W * \log_2(1 + \frac{(\frac{P_{i,k}^T}{D_{i,j}^\alpha})}{n*W})} * t \quad (5.2)$$

where W is a spectral width and n is spectral noise density. $D^{i,j}$ is the distance between the mobile terminal u_i and the base station b_j . From Equation (5.2), we can see that the energy increases with increase in the transmitting power. t is the communication time in seconds. For any wireless transmission the transmitting power gets attenuated as the distance between mobile terminal and the base station increases. α is the attenuation factor and depends upon the environmental condition and antenna characteristics. The relation between the transmission power and the distance is given by (Banerjee and Misra, 2002):

$$P_{i,k}^T = D_{i,j}^\alpha * \epsilon \quad (5.3)$$

where ϵ is the proportionality constant. If $P_{k,i}^R$ is the receiver power at the destination node u_k , the relation between the transmission and reception power is given by (Dong, Qunfeng and Banerjee, Suman and Adler, Micah and Misra, Archan, 2005),

$$P_{i,k}^T = \frac{P_{k,i}^R * D_{i,j}^\alpha}{\sigma} \quad (5.4)$$

where σ is the proportional constant.

5.3.2 Energy on Channel Scanning

When a mobile terminal is on an ongoing call, it is allocated a dedicated channel for the communication. At a given time, it can be associated with only one radio channel, unless it has an option for multi-radio communication. During an ongoing session, the mobile terminal continuously advertises the message and listens from available base stations in the vicinity, i.e., it continuously polls for available base stations. The polling process requires an enormous amount of energy as it involves continuous sending and receiving of messages. The mobile terminal collects the set of available base stations which ranges from 32 to 255 base stations. Whenever there is handover process initiated the mobile terminal scans these base stations. Scanning uses battery power and is not an energy efficient process, as the mobile terminal has to scan continuously. We propose to collect the information of the available base stations following the IEEE 802.21 standard. The standard defines a media-independent handover framework that can significantly reduce the complexity for handover between heterogeneous network technologies. The MIH protocol is located in the network and mobile terminal. It collects the information of the QoS and energy parameters for the ongoing session. Depending upon these parameters, decision of scanning available base stations is taken. With this scheme, the mobile terminal does not have to poll the available base stations instead the MIH gathers this information for the mobile terminal. The intelligent MIH connection monitoring manager sits between the application and device radio modem to monitor the wireless access network status and availability. Link manager is responsible for managing local link. It controls the local link by responding to MIH commands. The neighboring base station information can be collected using the Media-Independent Information Service (MIIS) in IEEE 802.21.

During the handover process the mobile terminal has to continuously scan for all the available channels, which involves prominent energy consumption. The energy consumption on channel scanning E_i^S depends on the time for which the channel is scanned and the number of channels scanned. The unit of E_i^S is joules. It increases with the increase in the number of channels and increase in the scanning interval, modeled as follows:

$$E_i^S = \sum_{j=1}^n E_{i,j}^{Scan} \quad (5.5)$$

where n is the number of channels and $E_{i,j}^{Scan}$ is energy required by u_i to scan channel j which is given by (Choi and Choi, 2007b),

$$E_{i,j}^{Scan} = (DIFS + \frac{aCW_{min}}{2} + minChannelTime) * P_{i,j}^{Scan} + (T_{p-req} - DIFS - \frac{aCW_{min}}{2}) * P_{i,j}^{Scan} \quad (5.6)$$

where $DIFS$ is the Distributed Inter Frame Space. The $aSlotTime$ is used to calculate the time interval between frames. The average $aSlotTime$ is $9\mu s$. The $minChannelTime$ controls the duration of the scanning and guarantees the reception of probe response frame. Typically, $minChannelTime$ should be larger than $DIFS + aCW_{min} * aSlotTime$, otherwise the user has to wait without receiving probe response after broadcasting probe request. If no probe response is received it is assumed that there is no access point available in this channel. T_{p-req} is the transmission time of a probe request. aCW_{min} is the initial value of contention window and is a physical layer parameter. P_{ij}^{Scan} is the power required by i^{th} channel to scan the j^{th} channel.

When a mobile device moves from one cell to another cell or from one network to another network, the call has to be transferred without dropping or degrading the quality. For a mobile device to maintain seamless connectivity, HDP plays an extremely distinguishable role.

5.4 Problem Formulation

In this section, the problem formulation is presented for minimizing the energy consumption at the mobile terminal through optimal handover while still guaranteeing the communication quality. The MOS is used as the main metric for measuring the communication quality. For the detail on how to compute MOS, please refer to our work in chapter 4 and paper (Jadhav, Zhang and Huang, 2012). As shown in Figure 5.2, consider an ongoing VoIP call connection between the mobile terminals u_i and u_k . When the MOS value of the current communication drops below a threshold \bar{M} , mobile terminal u_i initiates a handover and sends

a request to the handover controller. We will focus on the optimal based station selection algorithm performed at the handover controller with the objective to minimize the total energy consumption for data communication and channel scanning. The handover problem is formulated as the following optimization problem:

$$\text{minimize } E_{i,k}^{RT} + E_i^S \quad \text{subject to } M_{i,m} \geq \bar{M}; \quad (5.7)$$

where $M_{i,m}$ is the MOS value of the current connection, \bar{M} is the threshold MOS value.

5.5 Proposed Solution

In this section, we present the optimal and heuristic solutions for choosing the base station that minimizes the transmitting power while maintaining the call quality. We assume that the processing of the handover is done at the handover controller. The details on how to obtain these parameters will be described in the next section.

5.5.1 Optimal Solution

In this section, we present the solution for choosing the optimal base station during handover. A handover scenario in a heterogeneous network is considered, where several radio access technologies are available for users to handover the ongoing call session. Given a base station b_j and a mobile terminal u_i , let $M_{i,j}$ and $P_{i,j}$ denote the MOS and transmitting power between u_i and b_j , respectively. The mobile terminal u_i experiences poor call quality. Our goal is to design an efficient handover scheme to select the best base station b_m for u_i , by which the energy consumption is minimized, and maintaining the call quality.

From Equation (5.7), it can be seen that the energy increases with the increase in the transmitting power. For selecting the optimal base station b_m for the handover, we set threshold value for the MOS and the transmitting power. The MOS value and transmitting power of the current ongoing call are set as the threshold values M and P . If the base station b_m does not meet the quality criteria it is rejected for the further processing. If it meets the criteria, the

transmitting power of this base station is compared with the threshold value. This process is repeated for available base stations. If no base stations are found meeting this criteria the handover process is dropped. Let B_j be the set of base stations which satisfy the bandwidth requirement for the mobile terminal u_i . Let b^* be the optimal base station selected. The detailed algorithm for base station selection is given in Algorithm 2.

```

input :  $B = \{b_m : (M_{i,m}, P_{i,m})\}$ ,  $b_j$ ,  $u_i$ 
output: optimal  $b^*$ 

 $M = M_{i,j}$ ;  $P = P_{i,j}$ ;  $b^* = b_j$ ;
for each  $b_m : (M_{i,m}, P_{i,m})$  in  $B$  do
    if ( $M_{i,m} \leq M$ ) then
        | Reject the base station  $b_m$  ;
    else if ( $P_{i,m} \leq P$ ) then
        |  $M = M_{i,m}$ ;  $P = P_{i,m}$ ;  $b^* = b_m$ ;
    end
end

```

Algorithm 2: Optimal Base Station Selection

5.5.2 Heuristic Solution

In this section, we present a heuristic solution for choosing the base station during handover. We consider the handover scenario mentioned in section 5.1. Our goal is to design an efficient handover scheme to select the best base station b_m for the mobile terminal u_i , by which the energy consumption is minimized, and the quality of the call is maintained. The base station b_m is then selected for the handover by calculating the reward function for transmitting power and MOS, $R_{i,m}(P^T)$ and $R_{i,m}(M)$.

$$R_{i,m}(P^T) = \begin{cases} 0, & \text{when } P_{i,m}^T > P_{i,j}^T, \\ \frac{P_{i,j}^T - P_{i,m}^T}{P_{i,j}^T - (\min_{k=1}^n (P_{i,k}^T))}, & \text{when } P_{i,m}^T < P_{i,j}^T \end{cases} \quad (5.8)$$

where $\min_{k=1}^n (P_{i,k}^T)$ is the base station having minimum transmission power. $P_{i,j}^T$ is the transmission power of i^{th} mobile terminal with the current base station

and $P_{i,m}^T$ is the transmission power of i^{th} mobile terminal with the available base station. The MOS reward function is calculated as:

$$R_{i,m}(M) = \begin{cases} 0, & \text{when } M_{i,m} < M_{i,j}, \\ \frac{M_{i,j} - M_{i,m}}{M_{i,j} - (\max_{k=1}^n(M_{i,k}))}, & \text{when } M_{i,m} > M_{i,j} \end{cases} \quad (5.9)$$

where $\max_{k=1}^n(M_{i,k}^T)$ is the base station having maximum MOS. $M_{i,j}$ is the transmission power of the current base station and $M_{i,m}$ is the transmission power of the available base station. The reward function $R_{i,m}$ is calculated as (Munir and Gordon-Ross, 2009):

$$R_{i,m} = \omega R_{i,m}(P^T) + (1 - \omega) R_{i,m}(M) \quad (5.10)$$

where ω is the importance weighted factor. It can be assigned a value between 0.1 to 1 and is chosen based on the experience or historical data. In this study for simulation we have chosen $\omega = 0.1$. The base station with maximum reward has optimum energy and QoS. The base station with maximum reward is selected. Once the new base station is selected for the handover, the transmitting power of the mobile terminal is adjusted depending upon the benefit function.

Benefit Function: The RNC chooses a base station depending upon the value of the reward function. When a reward function is chosen, each reward is associated with a benefit function. In this work, we use transmitting power for the benefit function. Benefit functions gives how much transmitting power is saved or spent in extra by selecting the new base station. It helps us to know whether the new base station has more transmitting power than the current base station. Whenever an action is taken, i.e., when a handover is executed by selecting a new base station, the mobile terminal adjusts the transmitting power. This adjustment depends upon how much power is required for the mobile terminal to maintain the ongoing active link. If the energy required to continue the communication with the new base station increases, the transmission power level has to be increased. In this case the benefit function is negative. If energy required to continue the communication with the new base station decreases, then the transmission power level has to be decreased. In this case the benefit function is positive. The benefit function $b(i, m)$ is given by

$$b(i, m) = P_{i,j}^T - P_{i,m}^T \quad (5.11)$$

where $P_{i,m}^T$ is the transmitting power of mobile terminal u_i connecting to new base station b_n after handover. $P_{i,j}^T$ is the transmission power of the u_i before handover.

In the heuristic approach we propose a solution to select an energy efficient base station by calculating the reward function using the Equation (5.8, 5.9, 5.10 and 5.11). The base station with maximum reward is selected as it has optimum energy and QoS. Heuristic method is faster than optimal methods but it might not be optimal. The heuristic approach can be used where the speed of the process is important as it does not require exhaustive search and iterations as in optimal approach. The heuristic approach can lead to little inaccuracy as it based on experiential factors or based on the historical data.

In optimal approach we find the energy efficient base station by using Algorithm 2. Optimal solution at a times might be slower than heuristic as it involves systematically selecting value from the given set. It might involve few iterations before selecting the right value. An optimal approach finds the best value from the set for given objective function.

Figure 5.3 and Figure 5.4 show the transmitting power and MOS values for the optimal and heuristic approaches.

5.6 Handover Design Protocol

This section gives the details of our energy-based handover protocol design. As show in Figure 5.5 the handover process is divided into three steps:

- 1. Handover Initiation and Network Discovery:** When a mobile terminal detects its signal level going down, it initiates the handover process by sending a message to currently attached base station. Once the handover procedure is initiated, scanning of the available base station from the polling list starts. Scanning is performed when the mobile terminal joins the network or performs the handover. During the scanning operation, the current ongoing communication is paused for a time interval during which mobile terminal listens to the base

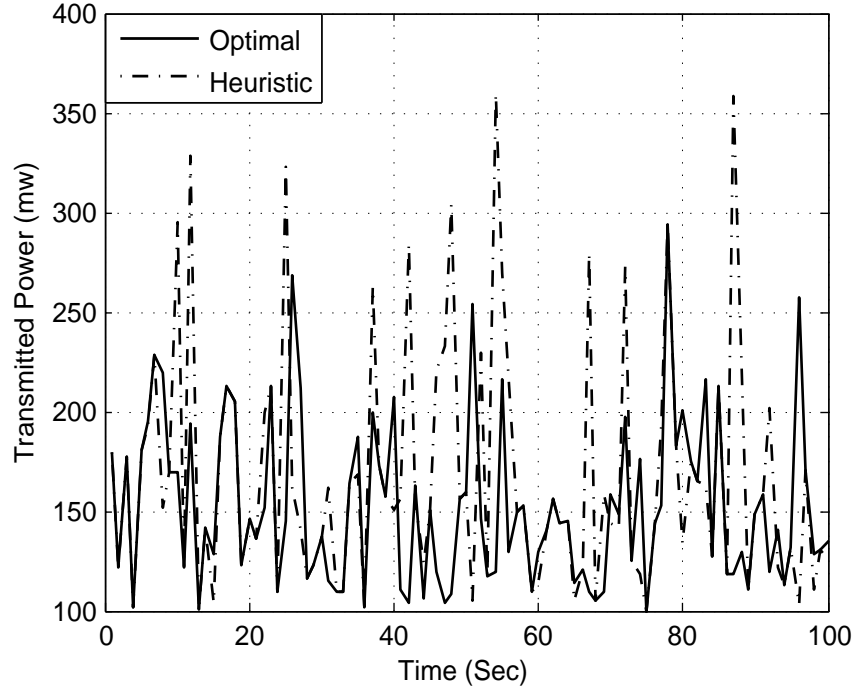


Figure 5.3: Transmitting Power for Optimal vs Heuristic

station. The mobile terminal is allocated a scan interval during which it listens to the base station frequency. It gets information about the link parameters, this is called ranging process. When the mobile terminal joins the network it performs initial network entry procedure, where it scans the available base stations by listening to the base station frequency. During the handover, the information of the neighboring base stations is collected by the MIH protocol. Once it receives the set of base stations, it scans each of the base stations to calculate the QoS parameters and the energy required to maintain the ongoing communication.

During the scanning operation, the current ongoing communication is paused for a time interval during which mobile terminal listens to the base station. The mobile terminal is allocated a scan interval during which it listens to the base station frequency. It gets information about the link parameters, this is called ranging process.

2. Handover Decision: The results of the scanning process are used for the handover decision process. During the decision process, new base station is selected for the mobile terminal based on the selection criteria. QoS parameter and transmission power are used as selection criteria. Mobile terminal sends this information to the handover controller, where the controller selects the base station

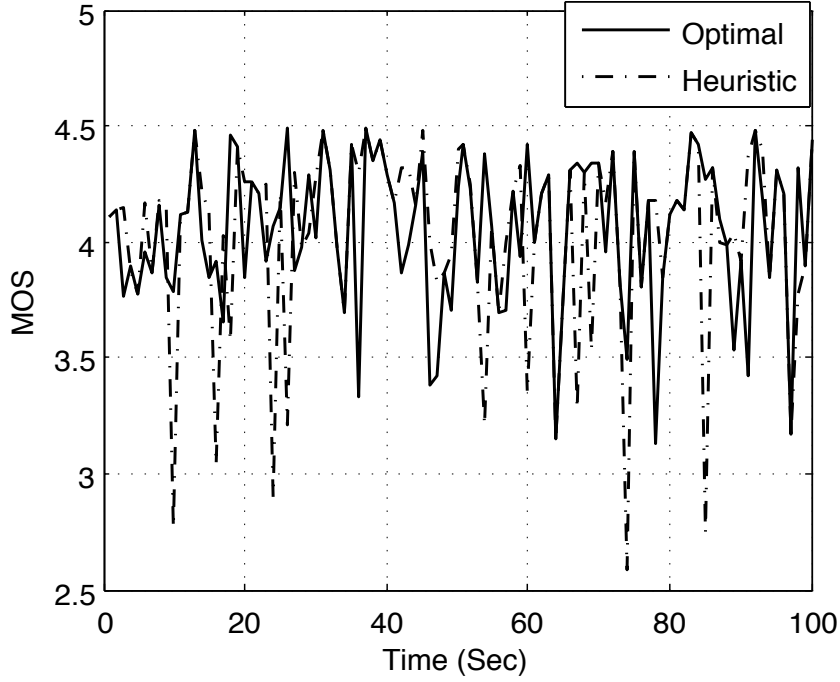


Figure 5.4: MOS for Optimal vs Heuristic

using the above algorithm. From Equation (5.2), it can be seen that the energy consumption is related to the transmission power. It can be seen that the energy increases with the increase in the transmission power. The transmission power of the mobile terminal is determined primarily by factors such as, link condition, delay and packet loss rate. It is also determined by the distance between the mobile terminal and base station, which is called as link distance. Our objective is to choose the link which minimizes the energy consumption. This objective is achieved by selecting a link with reliable and efficient transmission power.

3. Handover Execution: Once the base station is selected, handover execution takes place along with the authentication, authorization and registration is done, communication is transferred to the new base station. If the handover is successful, i.e., new base station is selected for the handover, the controller sends the transmission power required to operate the new link. The mobile terminal then adjusts the transmission power in order to maintain the successful transmission. It is assumed that the handover controller has knowledge of the QoS parameters of each base station and the transmission power for each wireless link between the mobile terminal and base stations.

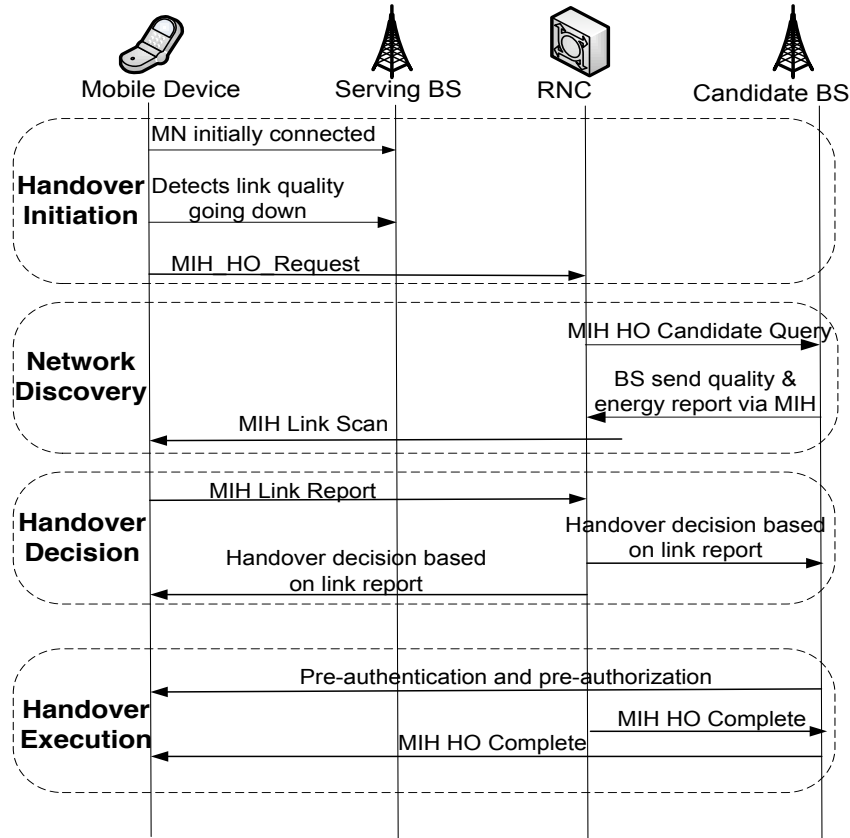


Figure 5.5: Handover Design Protocol

5.7 Simulation Setup

To evaluate the proposed scheme, energy-based handover protocol is implemented in MATLAB. It is simulated in an integrated environment with WiMAX and UMTS networks, and its performance is compared with the RSS-based handover (Sarddar et al., 2010). In this study, transmission power, MOS and HDP are measured.

In our simulation, 10 UMTS base stations and 4 WiMAX base stations are deployed in a $10000m \times 10000m$ area. The mobile terminals are uniformly placed in the UMTS or WiMAX cells. Each of the UMTS or WiMAX cell has a base station. All the base stations of the UMTS network are connected to the RNC and all the base stations of the WiMAX network are connected to ASN which in turn is connected to RNC. Handovers between UMTS and WiMAX networks are managed by the RNC. The diameter of a UMTS cell is configured to 2 km, and the diameter of a WiMAX cell is configured to 3 km. In case of the UMTS

network when a mobile terminal makes a VoIP call, the voice packets are carried from mobile terminal to the RNC through Node-B. For the WiMAX network when a mobile terminal makes a VoIP call, the voice packets are carried from mobile terminal to the RNC through ASN.

In our simulation of energy-based handover protocol, a 2D random walk model is used to simulate the movement of the mobile terminal. Random Walk mobility model is proposed to mimic their movement behavior. The random walk model is a stateless mobility process, where the decision of future is not taken based on the information about the previous state. That is, the current parameter information is independent with its previous parameter information. Every mobile user moves one step ahead or one step behind or stays at same position randomly. When the mobile user moves away from the base station the packet loss and delay changes. The change in the packet loss and the delay is computed according. The movement of every mobile terminal is associated with 1. the moving direction of the mobile terminal 2. step size. Because some mobile terminals are believed to move in an unexpected way, random walk mobility model is proposed to mimic their movement behavior (Bai and Helmy, 2004). Each time the mobile terminal walks for the distance L , a new direction is randomly chosen (i.e. four possibilities: 1) forward, 2) backward, c) left and d) right) then walks for another distance L . We compare our energy-based handover protocol with RSS-based handover protocol proposed in (Grønsund, Grøndalen, Breivik, and Engelstad, 2007). The RSS is calculated using the following function

$$RSS = -62.5 - 26.5 * \log_{10}(d) \quad (5.12)$$

where d is the distance between a mobile terminal and a base station. If mobile terminal detects the RSS value below the threshold, the mobile terminal scans for the available networks and handovers the call to the base station providing higher RSS. If the mobile terminal fails to find a better base station, the handover request is rejected, and the mobile terminal continues the call by connecting to the same base station.

5.8 Performance Metrics

In our simulations, we use the following three metrics to evaluate the performance of WiMAX and UMTS in terms of energy consumption and end-to-end QoS for VoIP.

5.8.1 Transmitting Power

Wireless signal transmission is sending the signal from one device to other by magnetic induction, electromagnetic radiation or microwave transmission. Each mobile terminal has an antenna which is used for transmission and reception of the radio signal. The transmission power of the mobile terminal can be either fixed or variable. In case of fixed transmission power, all the terminals in the network operate at a fixed power irrespective of the distance, type of application or type of antenna. Fixed transmission is not an efficient method, as the mobile terminal away from the base station need more transmission power and those which are near need less transmission power. Since fixed transmission operate at a single transmission power, it results in poor QoS for the far mobile terminals. In our scheme we use variable transmission power, where the transmission power of a mobile terminal can be adjustable dynamically and is controlled by a power control system. The transmission power is calculated using Equation 5.3 The mobile terminal which is far from the base station can use higher transmission power and thus can maintain the QoS. However, there is a trade-off between then the transmission power and energy consumption. Higher the transmission power, higher is the energy consumption. Hence the optimum transmission power should be selected in order to maintain the balance between the QoS and energy consumption.

5.8.2 MOS

The detailed explanation of MOS is given in the Chapter 4.

5.8.3 Handover Dropping Probability

If a handover request is not processed, the handover is dropped. The corresponding probability is called as Handover Dropping Probability (HDP). The

call still continues with the current attached base station. We consider the handover dropping probability as a constraint for handover. If the handover dropping probability is high, this means there are more chances of handover processes being dropped, which causes more energy consumption as mobile terminal has to initiate the handover process again.

5.9 Simulation Results

In this section, we compare the performance of energy-based and RSS-based handover schemes. We use the parameters presented in the Section 5.8 for the evaluation.

5.9.1 Transmitting Power

Figure 5.6 shows the transmitting power values for the proposed optimal energy-based and RSS-based handover schemes (Grønsund, Grøndalen, Breivik, and Engelstad, 2007).

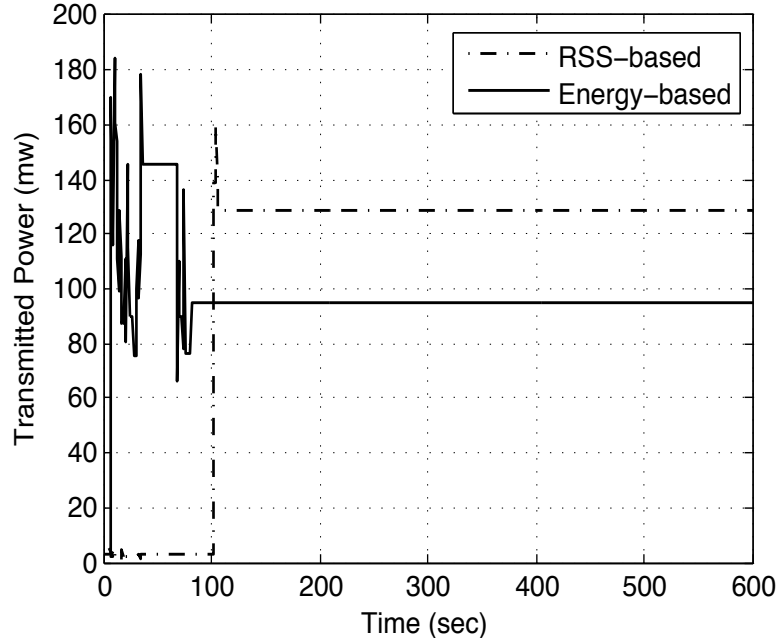


Figure 5.6: Transmission Power

In case of energy-based handover scheme initially, the mobile terminal has a transmitting power value of 180mw. Since the mobile terminal does not have

a base station with better transmission power, it continues the service with the current base station until it finds a suitable base station at approximately the 20sec. Once the mobile terminal is able to get a base station with a lower transmission power value, it performs a handover. Until the end of the first minute, the mobile terminal maintains the connection with the current serving base station. The mobile terminal detects the decline in quality after the first minute and performs a search operation looking for the base station with lower transmitting power values among the available base stations for handover. In the case of the RSS-based scheme, the mobile terminal is initially transmitting at low power until 100ms, where the handover occurs. During the handover the mobile terminal selects the base station having a higher RSS. Since the transmission power is not considered in RSS-based handover, the mobile terminal ends up in selecting base station with a higher transmission power, thus having higher energy consumption. From the graph, we can see that in energy-based scheme transmitting power of the call is maintained and the mobile terminal tends to select the base station which provides less transmission power values every time it decides to make a handover.

5.9.2 MOS

Figure 5.7 shows the MOS values for the proposed energy-based and RSS-based handover scheme.

Initially, the mobile terminal has a MOS value of 3.5. Until the end of the first minute, the mobile terminal maintains the connection with the current serving base station. The mobile terminal detects the decline in quality after the first minute and performs a search operation looking for the base station with highest MOS. Once the mobile terminal is able to get a base station with MOS value higher than the threshold, it performs a handover. From the graph, we can see that in energy-based handover scheme MOS value of the call is maintained 3.5, whereas in RSS-based scheme MOS value drops to 1.9.

5.9.3 Handover Dropping Probability

Figure 5.8 shows the HDP of RSS-based handover is higher than the MOS-based handover schemes. The handover process fails for various reasons such as, no

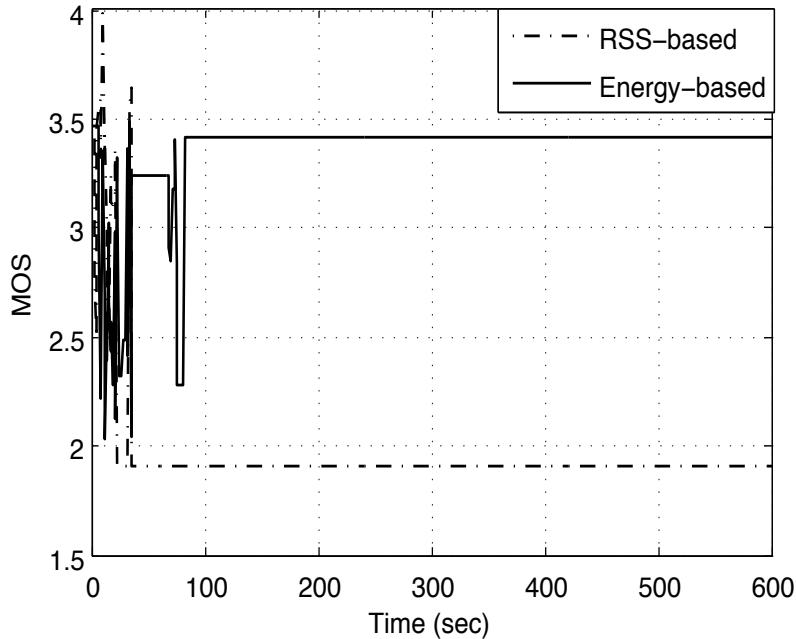


Figure 5.7: MOS

sufficient bandwidth or no enough wireless resources to provide strong signal strength to the mobile terminal receivers.

5.10 Discussion

NGWN consists of multiple networks and can support rich multimedia applications which require high transmitting power on the mobile terminal. In this regard, battery power of the mobile terminal to support multimedia applications need to be stronger. In view of this, we propose a new energy efficient handover algorithm to conserve the energy of mobile terminal with a constraint of considerable QoS to support the multimedia applications. Simulation results show a substantial improvement in terms of call dropping probability, power consumption and MOS when compared to the traditional RSS handover technique.

With our scheme mobile terminal does not have to scan the available base stations, MIH gathers this information for the mobile terminal. The neighboring base stations information can be collected using the Media-Independent Information Service (MIIS) in IEEE 802.21. The intelligent MIH connection monitoring manager sits between the application and device radio modem to monitor the wireless access, network status and availability. Link manager is responsible for

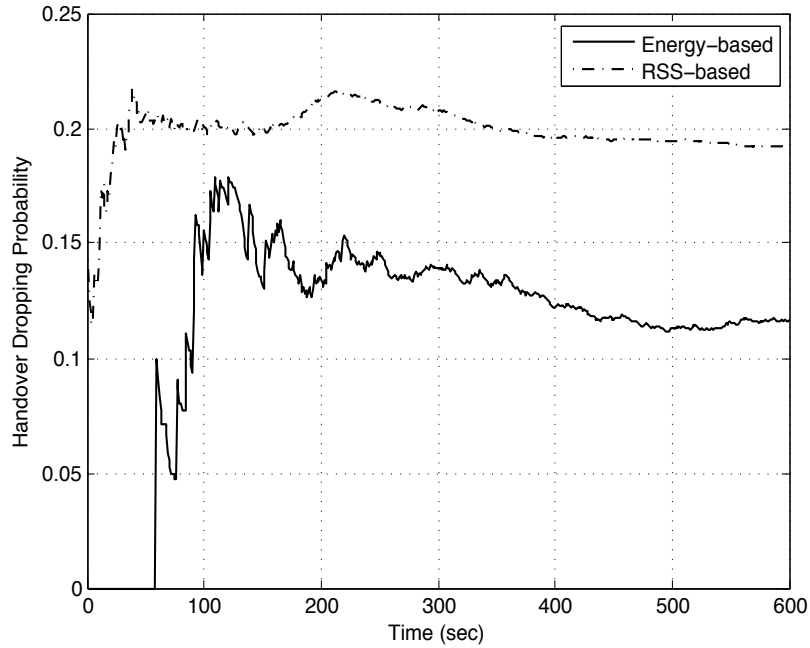


Figure 5.8: Handover Dropping Probability

managing local link. It controls the local link by responding to MIH commands. This work investigates the problem of minimizing the energy consumption at the handset side through optimal handover with guarantee on the communication quality. We propose both a centralized optimal solution and a heuristic solution for computing the optimal base station for handover, and design an energy-efficient handover scheme based on IEEE 802.21 for next generation wireless networks. The numerical results show that our scheme is energy efficient and provides better quality compared to the traditional RSS-based scheme.

To conserve the battery power in order to support more applications and support for a longer time, the mobile terminal need an energy efficient handover process. In this paper, the problem of energy efficiency has been investigated in detail.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

The deployment of the NGWNs has posed great challenges due to the demanding requirements to integrate different radio access technologies and to provide high QoS for applications such as VoIP traffic. NGWNs focus on convergence of different radio access technologies providing good QoS for applications such as VoIP and video streaming. The voice applications over IP networks are growing rapidly due to their increasing popularity. To meet the demand of providing high-quality of VoIP it is imperative to design suitable QoS model. The NGWNs are expected to be all-IP networks with high data rates, optimized QoS and seamless connectivity. To achieve these challenges, handover plays a very important role in maintaining the connectivity. During the process of handover, energy consumption is a major problem for battery operated devices. These devices have very limited power and are expected to operate for longer time. In order to overcome these challenges we carried out an extensive study to evaluate the performance of leading wireless technologies. We also devised a QoS-based energy efficient handover scheme for NGWNs.

In Chapter 3, we investigated the QoS performance of WiMAX and UMTS for supporting VoIP applications since UMTS and WiMAX are emerging as the promising technologies for building the NGWNs. Simulation study was conducted to evaluate the QoS performance of WiMAX and UMTS for supporting VoIP traffic. Simulation modules were designed in OPNET for WiMAX and UMTS, and extensive simulations were carried out to evaluate and analyze sev-

eral important performance metrics such as MOS, end-to-end delay, jitter and packet delay variation. Simulation results show that WiMAX outscores UMTS with a sufficient margin, and is the better technology to support VoIP applications compared with UMTS. This study is our first step towards exploring possible implementations of the next generation wireless networks.

In Chapter 4, we proposed a novel scheme compliant with the IEEE 802.21 standard for handover in an integrated scenario with UMTS and WiMAX networks. NGWNs are expected to provide high data rates and optimized quality of service to multimedia and real-time applications over the IP networks. To achieve these goals, handover plays a very critical role in maintaining the seamless connectivity when mobile terminals move across different cells or networks. The call quality, measured using MOS, is used as the major metric for handover optimization. The proposed MOS-based handover scheme is compared with the traditional RSS-based handover scheme. The numerical results demonstrate that the proposed scheme can maintain high call quality and reduce the probabilities for both handover dropping and call dropping.

Energy consumption has been a major issue on battery powered handsets like cellphones due to the expectation to make the devices operate as long as possible. Handover, the process of transferring an ongoing call from one base station to another to guarantee seamless connection, has a significant impact on energy at the handsets. Chapter 5 investigates the problem of minimizing the energy consumption at the handset side through optimal handover with guarantee on the communication quality. Both a centralized optimal solution and a heuristic solution are proposed for computing an optimal base station for handover, and designed an energy-efficient handover scheme based on IEEE 802.21 for next generation wireless networks. The simulation results show that our scheme is energy efficient and provides better quality compared with the traditional RSS based scheme.

In summary, this thesis examined the QoS models of UMTS and WiMAX networks. It provides efficient solutions for making an efficient quality-based handover in homogeneous and heterogeneous networks. The proposed scheme in this thesis is energy efficient and will help to conserve energy at the mobile terminals.

6.2 Future Work

In the future, when UMTS and WiMAX allow users to select any of the available networks, the classifications of QoS for different services in different networks provide user an easy way switch to the network which is best suited for the real-time applications. In this regard, we will extend our work to classify the QoS and make it transparent to the user. This classification will not only help the user but also the service providers. This classification will allow the user to select the best available network depending upon the user requirements. It will provide the guaranteed QoS requirement during network switching. For the network providers, this work will help reduce the deployment cost by using other network infrastructure to maintain the call continuity.

Energy consumption is the hottest topic lately in the research community. Network infrastructure consists of many components such as, servers, access network components such as base stations, radio towers etc. Particularly the access network components consume a lot of energy as they are the front end of the network supporting millions of users. Base stations need to be active all the time and process requests of many users simultaneously. This procedure makes the base station vulnerable to high energy consumption. In this regard, we will continue our energy related models to achieve greater energy efficiency in base stations. Especially, in the future, base stations will be processing more requests than today due to the increased number of users every day. In addition, the same work can be extended to femto-cells (i.e., small cells) which are considered as fast growing elements in the network. Small cells provide capacity coverage and data offload for users which are not subsequently under coverage of its main base station.

Appendix A

Description of Parameters Used in Chapter 3

Parameter	Description
M	Mean Opinion Score
R	R-score
I_s	impairment caused by signal-to-noise
I_e	equipment impairment factor
I_d	impairment caused by sum of all the delays
A	Advantage factor
D_n	delay in the network
D_e	delay caused by encoder
D_d	delay caused by decoder
D_c	delay caused by codec
$t(i)$	packet transmitted time at the transmitter
$t'(i)$	packet received time at the receiver

Appendix B

Description of Parameters Used in Chapter 4

Parameter	Description
b_i	i^{th} base station
m_j	j^{th} mobile terminal
R_o	effect of noise
I_s	impairment caused by signal-to-noise
I_e	equipment impairment factor
I_d	impairment caused by sum of all the delays
A	Advantage factor
d_{codec}	delay caused by coder and encoder
d_{payout}	play-out delay which is caused due to buffering
$d_{network}$	delay in the network
$e_{network}$	rate of loss of packets in the network
e_{payout}	rate of loss packets during the payout of voice signal
e	Network and play-out buffer losses
d	total delay composed of three components codec delay, play-out delay and network delay
B	set of UMTS and WiMAX base stations deployed in urban area

M	set of mobile terminal
b_i	i^{th} base station
m_j	j^{th} mobile terminal
$p_{i,j}$	packet loss rate between b_i and m_j
$d_{i,j}$	average packet delivery delay between b_i and m_j
R	R-Score
$M_{i,j}$	MOS value for the VoIP call between base station b_i and mobile terminal m_j
c_i	available bandwidth capacity at the base station b_i .
B_j	set of base stations which satisfy bandwidth requirement for the mobile terminal m_j

Appendix C

Description of Parameters Used in Chapter 5

Parameter	Description
B	set of UMTS and WiMAX base stations deployed in urban area
U	set of mobile terminal
b_i	i^{th} base station
u_j	j^{th} mobile terminal
u_k	k^{th} mobile terminal
$P_{i,k}^C$	average circuit power of the mobile terminal to operate on the i^{th} communication link
$E_{i,k}^{L,tot}$	total energy consumed by the mobile terminal
$r_{i,k}$	the transmission rate (bit/second)
t	the communication time (second)
$E(P_{i,k}^c)$	energy communication circuit used during the active communication
$E_{i,k}^{RT}$	energy required for reliable transmission over link (i, k)
W	spectral width
$D_{i,j}^\alpha$	distance between mobile terminal u_i and the base station b_j
n	spectral noise density
$P_{i,k}^T$	transmission power of the u_i
α	attenuation factor

$P_{k,i}^R$	receiver power at the destination node u_k
σ	proportional constant
n	j^{th} number of channels
t_s	scan time on each channel
$E_{i,j}^S$	energy required by u_i to scan channel j
$T_{p.req}$	transmission time of a probe request
aCW_{min}	initial value of contention window and is a physical layer parameter
P_{ij}^{Scan}	power required by i^{th} channel to scan the j^{th} channel
$M_{i,m}$	MOS value of the current connection
$\bar{t}_{i,j}$	expected time for communication over channel (i, j)
b_m	optimal base station for the handover
B_j	set of base stations which satisfy bandwidth requirement for the mobile terminal m_j
$R_{i,m}(P^T)$	reward function for transmitted power
$R_{i,m}(M)$	reward function for MOS
$min(P_{i,n}^T)$	base station having minimum transmission power
$max(M_{i,n}^T)$	the base station having maximum MOS
ω	the weighted factor
λ_n	new connection request rate
λ_h	handover request rate
μ_n	new call
μ_h	handover call departure
$P_{i,j}^{hd}$	handover dropping probability on channel (i, j)

References

- Liang, Q. (2002). Bandwidth utilization and signal strength-based handover initiation in mobile multimedia cellular networks. *IEEE International Symposium Circuits and Systems, ISCAS 2002*, 3, 21-24.
- Liao, H. and Tie, L. and Du, Z. (2006). A vertical handover decision algorithm based on fuzzy control theory. *First International Multi-Symposiums Computer and Computational Sciences, IMSCCS'06*, 2, 309-313.
- Dai, Z. and Fracchia, R. and Gosteau, J. and Pellati, P. and Vivier, G. (2008). Vertical handover criteria and algorithm in IEEE802. 11 and 802.16 hybrid networks. *International Conference on Communications, ICC'08*, 2480-2484.
- Desset, C. and Ahmed, N. and Dejonghe, A. (2009). Energy savings for wireless terminals through smart vertical handover. *IEEE International Conference on Communications, ICC'09*, 1-5.
- Ma, L. and Yu, F. and Leung, V.C.M. and Randhawa, T. (2004). A new method to support UMTS/WLAN vertical handover using SCTP. *IEEE International Conference on Wireless Communications, 2004*, 11(4), 44-51.
- Choi, H.H. and Cho, D.H. (2005). Takeover: a new vertical handover concept for next-generation heterogeneous networks. *61st IEEE International Vehicular Technology Conference, VTC 2005-Spring*, 4, 2225-2229.
- Inzerilli, T. and Vegni, AM and Neri, A. and Cusani, R.(2008). A location-based vertical handover algorithm for limitation of the ping-pong effect. *IEEE International Conference on Wireless and Mobile Computing, WIMOB'08*, 385-389.
- Lin, H.P. and Juang, R.T. and Lin, D.B.(2005). Validation of an improved location-based handover algorithm using GSM measurement data. *IEEE Transactions on Mobile Computing, 2005*, 4(5), 530-536.

- Uysal-Biyikoglu, E., Prabhakar, B., and El Gamal, A. (2002). Energy-efficient packet transmission over a wireless link. *IEEE/ACM Transactions on Networking (TON)*, 2002, 10(4), 487–499.
- Jadhav, S. and Zhang, H. and Huang, Z.(2012). MOS-based Handover Protocol for Next Generation Wireless Networks. *IEEE 26th International Conference on Advanced Information Networking and Applications (AINA)*, 2012, 479–486.
- Fernandez, J.C. and Taleb, T. and Guizani, M. and Kato, N.(2009). Bandwidth aggregation-aware dynamic QoS negotiation for real-time video streaming in next-generation wireless networks. *IEEE Transactions on Multimedia*, 2009, 11(6), 1082–1093.
- Fernandez, J.C. and Taleb, T. and Guizani, M. and Kato, N.(2010). LTE-advanced: next-generation wireless broadband technology [Invited Paper]. *IEEE Transactions on Wireless Communications*, 2010, 17(3), 10–22.
- Joung, J. and Song, J. and Lee, S.S.(2008). Flow-based QoS management architectures for the next generation network. *Electronics and Telecommunication Research (ETRI) journal*, 2008, 30(2), 238.
- Niyato, D. and Hossain, E.(2009). Cognitive radio for next-generation wireless networks: an approach to opportunistic channel selection in IEEE 802.11-based wireless mesh. *IEEE Transactions on Wireless Communications*, 2009, 16(1), 46–54.
- Kim, R.Y. and Mohanty, S.(2010). Advanced power management techniques in next-generation wireless networks. *IEEE Communications Magazine*, 2010, 48(5), 94–102.
- Han, C. and Harrold, T. (2011). Green radio: radio techniques to enable energy-efficient wireless networks. *IEEE Communications Magazine*, 2011, 49(6), 46–54.
- Salsano, S. and Polidoro, A. and Mingardi, C. and Niccolini, S. and Veltri, L. (2008). SIP-based mobility management in next generation networks. *IEEE Journal on Wireless Communications*, 2008, 45(2), 92–991

- Yan, X. and Ahmet Şekercioğlu, Y. and Narayanan, S. (2010). A survey of vertical handover decision algorithms in Fourth Generation heterogeneous wireless networks. *Journal on Computer Networks, Elsevier, 2010, 54*(11), 1848–1863
- Vegni, A.M. and Tamea, G. and Inzerilli, T. and Cusani, R. (2009). A combined vertical handover decision metric for QoS enhancement in next generation networks. *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WIMOB 2009, 233–238*
- Saboji, SV and Akki, CB. (2011). Available Bandwidth Estimation in 4G Wireless Networks Using Mobile Agents. *International Journal of Research and Reviews in Information Sciences (IJRRIS), 2011, 1*(2).
- Feeney, L.M. and Nilsson, M. (2001). Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. *IEEE Twentieth Annual Joint Conference of Computer and Communications Societies, INFOCOM 2001, 3*, 1548–1557
- Edler, T. and Lundberg, S. (2004). Energy efficiency enhancements in radio access networks. *Ericsson Review, 2004, 81*(1), 42–51
- Jones, C.E. and Sivalingam, K.M. and Agrawal, P. and Chen, J.C. (2001). A survey of energy efficient network protocols for wireless networks. *wireless networks, Kluwer Academic Publishers, 2001, 7*(4), 343–358
- Claussen, H. and Ho, L.T.W. and Pivit, F. (2008). Effects of joint macrocell and residential picocell deployment on the network energy efficiency. *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008, 1–6*
- Feder, P. and Isukapalli, R. and Mizikovsky, S. (2009). WiMAX-EVDO interworking using mobile IP. *IEEE Transactions on Communications Magazine, 2009, 47*(6), 122–131
- Arjona, A. and Westphal, C. and Yla-Jaaski, A. and Kristensson, M. (2008) Towards high quality VoIP in 3G networks-an empirical study. *Fourth Advanced International Conference on Telecommunications AICT'08, 143-150*
- WiMAX QoS Classes to support Voice, Video, and Data Traffic,
http://www.tranzeo.com/allowed/Tranzeo_WiMAX_QoS_Classes_Whitepaper.pdf.

- The Real WiMAX Advantage,
http://www.visantstrategies.com/analysis/wimax_advantage.html.
- IEEE Standard for Local and metropolitan area networks - Part 21: Media Independent Handover Services. *IEEE 802.21-2008*.
- Opnet Tool Information, www.opnet.com.
- E-model (2000). The E-model, a Computational Model for Use in Transmission Planning. ITU-T Recommendation G.107.
- Zahid Ghadialy (2007). http://www.3g4g.co.uk/Tutorial/ZG/zg_pdp.
- 3GPP, T. (2005). 3GPP, Technical Specification Group Radio Access Network UTRAN Overall Description(Release 99).
- 3GPP, T. (2006). 3GPP, Technical SpeciGroup RAN, User Equipment (UE) Radio Transmission and Reception (FDD).
- Abichar, Z., Peng, Y., and Chang, J. (2006). WiMAX: The emergence of wireless broadband. *IT professional*, 8(4), 44–48.
- Adhicandra, I. (2010). Measuring Data and VoIP Traffic in WiMAX Networks. *Journal of Telecommunication*, 2010, 2.
- Agharebparast, F. and Leung, V. (2002). A framework for QoS support in the UMTS/GPRS backbone network using DiffServ. *Computing and informatics*, 2002, 21(2), 97–118.
- Akyildiz, I., Xie, J., and Mohanty, S. (2004). A survey of mobility management in next-generation all-IP-based wireless systems. *IEEE Wireless Communications*, 2004, 11(4), 16–28.
- Alamri, N. and Akkari, N. (2012). UMTS-WiMAX Vertical Handover in Next Generation Wireless Networks. *Arxiv preprint arXiv:1201.1038*.
- Andrews, J., Ghosh, A., and Muhamed, R. (2007). *Fundamentals of WiMAX*. Prentice Hall.
- Bai, F. and Helmy, A. (2004). A survey of mobility models. *Wireless Adhoc Networks. University of Southern California, USA, 206*.

- Balan, V., Eggert, L., Niccolini, S., and Brunner, M. (2007). An experimental evaluation of voice quality over the Datagram Congestion Control Protocol. *26th IEEE International Conference on Computer Communications (INFOCOM), 2007*, 2009–2017.
- Balasubramanian, N., Balasubramanian, A., and Venkataramani, A. (2009). Energy consumption in mobile phones: a measurement study and implications for network applications. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference, 2009*, 280–293. ACM.
- Banerjee, S. and Misra, A. (2002). Minimum energy paths for reliable communication in multi-hop wireless networks. In *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing, 2002*, 146–156. ACM.
- Belghith, A. and Nuaymi, L. (2008). Comparison of WiMAX scheduling algorithms and proposals for the rtPS QoS class. In *IEEE Wireless Conference, 2008. EW 2008. 14th European*, 1–6.
- Belghith, A., Nuaymi, L., and Maillé, P. (2008). Pricing of differentiated-QoS services WiMAX networks. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008.*, 1–6. .
- Bella, M., Olivier, M., and Eloff, J. (2005). A fraud detection model for Next-Generation Networks. In *Proceedings of the 8th Southern African Telecommunications Networks and Applications Conference (SATNAC 2005), Central Drakensberg, KwaZulu-Natal, South Africa*.
- Bin, L., Martins, P., Bertin, P., and Samhat, A. Inter-RAT Handover between UMTS and WiMAX, 2010.
- Black, U. (1999). *Voice over IP*. Prentice-Hall, Inc.
- Bolla, R., Bruschi, R., Davoli, F., and Ranieri, A. (2009). Energy-aware performance optimization for next-generation green network equipment. In *Proceedings of the 2nd ACM SIGCOMM workshop on Programmable routers for extensible services of tomorrow, 2009*, 49–54. ACM.

- Braga, A., Rodrigues, E., and Cavalcanti, F. (2006). Packet scheduling for VoIP over HSDPA in mixed traffic scenarios. In *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications, 2006*, 1–5. .
- Chakraborty, M. and Bhattacharyya, D. (2010). Overview of End-to-End WiMAX Network Architecture. *WiMAX Security and Quality of Service: An End-to-End Perspective*.
- Chang, B. and Chen, J. (2008). Cross-layer-based adaptive vertical handoff with predictive RSS in heterogeneous wireless networks. *IEEE Transactions on Vehicular Technology, 2008*, 57(6), 3679–3692.
- Chang, B., Chen, J., Hsieh, C., and Liang, Y. (2009). Markov decision process-based adaptive vertical handoff with RSS prediction in heterogeneous wireless networks. In *IEEE Wireless Communications and Networking Conference, 2009. WCNC 2009.*, 1–6.
- Chen, X., Wang, C., Xuan, D., Li, Z., Min, Y., and Zhao, W. (2003). Survey on QoS management of VoIP. In *IEEE International Conference on Computer Networks and Mobile Computing, 2003. ICCNMC 2003*, 69–77.
- Banerjee, Suman and Misra, Archan (2004). Power adaptation based optimization for energy efficient reliable wireless paths. In *Journal of Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks, 2004. Springer*, 854–865.
- Chen-Feng, W., Liang-Teh, L., and Der-Fu, T. (2011). A Novel Call Admission Control Policy Using Mobility Prediction and Throttle Mechanism for Supporting QoS in Wireless Cellular Networks. *Journal of Control Science and Engineering, 2011, 2011*.
- Choi, Y. and Choi, S. (2007a). Service charge and energy-aware vertical handoff in integrated IEEE 802.16 e/802.11 networks. In *26th IEEE International Conference on Computer Communications, INFOCOM 2007*, 589–597. IEEE.
- Choi, Y. and Choi, S. (2007b). Service charge and energy-aware vertical handoff in integrated IEEE 802.16 e/802.11 networks. In *26th IEEE International Conference on Computer Communications, INFOCOM 2007*, 589–597.

- Cicconetti, C., Eklund, C., Lenzini, L., and Mingozzi, E. (2006). Quality of Service Support in IEEE 802.16 Networks. *IEEE Network Magazine*, 2006, 20(2).
- Cole, R. and Rosenbluth, J. (2001). Voice over IP performance monitoring. *ACM SIGCOMM Computer Communication Review*, 2001, 31(2), 9–24.
- Demichelis, C. and Chimento, P. (2002). IP packet delay variation metric for IP performance metrics (IPPM). Technical report, RFC 3393, November.
- DESANTIS, M. (2006). Understanding Voice over Internet Protocol (VoIP). *IUS-CERT*.
- Ding, L. and Goubran, R. (2003). Speech quality prediction in VoIP using the extended E-model. In *IEEE Global Telecommunications Conference, 2003 (GLOBECOM'03)*, Volume 7, 3974–3978.
- Duffield, N. and Lo Presti, F. (2000). Multicast inference of packet delay variance at interior network links. In *Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2000*, Volume 3, 1351–1360.
- Ericsson (2004). Categorising telecommunications fraud an introduction for those new to the subject.
- ETSI, GSM (1998). General packet radio service; service description; stage 2.
- Garg, V. and Yu, O. (2000). Integrated QoS support in 3G UMTS networks. In *IEEE Wireless Communications and Networking Conference, WCNC 2000*, Volume 3, 1187–1192.
- Ghosh, A., Wolter, D., Andrews, J., and Chen, R. (2005). Broadband Wireless Access with WiMAX/802.16: Current Performance Benchmarks and Future Potential. *IEEE communications magazine*, 2005, 43(2), 129.
- Grønsund, P., Grøndalen, O., Breivik, T., and Engelstad, P. Fixed WiMAX Field Trial Measurements and the derivation of a Path Loss Model. *University of Oslo*.
- Guowang Miao, Nageen Himayat, Y. G. L. and Swami., A. Cross-Layer Optimization for Energy-EfWireless Communications: A Survey.

- Hallock, J. (2004). A Brief History of VoIP. *Evolution*.
- Holma, H. and Toskala, A. (2006a). *Hsdpa/Hsupa For Umts*. Wiley Online Library.
- Holma, H. and Toskala, A. (Eds.) (2006b). *HSDPA/HSUPA for UMTS* (First ed.).
- Holma, H., Toskala, A., (2000). *Wcdma for Umts*, Volume 4. Wiley Online Library.
- Hwang, I., Hwang, B., Ku, L., and Chang, P. (2008). Adaptive bandwidth management and reservation scheme in heterogeneous wireless networks. In *IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, SUTC'08*, 338–342.
- Institute InfoSec Reading Room, S. (2004). Voice Over Internet Protocol (VoIP) and Security.
- ITU-T, R. (1996). Methods for subjective determination of transmission quality. *International Telecommunication Union-Telecommunication Standardisation Sector (ITU-T)*.
- JDSU (2003). White Paper - VoIP Overview.
- Khan, F. (2008). A Generic Technique for Voice over Internet Protocol (VoIP) Traffic Detection. *IJCSNS*, 8(2), 52.
- Khan, M., Ismail, M., and Dimyati, K. (2010). Interworking between WiMAX and UMTS to provide seamless services. *Int. J. Phys. Sci*, 5(6), 852–864.
- Kim, D., Cai, H., Na, M., and Choi, S. (2008). Performance measurement over Mobile WiMAX/IEEE 802.16 e network. In *International Symposium on a World of Wireless, Mobile and Multimedia Networks, WoWMoM 2008*, 1–8. IEEE.
- Kim, Y. (2006). VoIP Service on HSDPA in Mixed Traffic Scenarios.
- Korotky, S. and Pfeiffer, T. (2009). Continuing advances in next-generation communication technologies, services, and networks. *Bell Labs Technical Journal*, 2009, 14(1), 1–5.

- Kunarak, S. and Suleesathira, R. Predictive RSS with fuzzy logic based vertical handoff algorithm in heterogeneous wireless networks. In *IEEE International Symposium on Communications and Information Technologies (ISCIT), 2010*, 1235–1240.
- Lee, C. and Morita, N. (2006). 1–15.
- Lee, H., Kwon, T., Cho, D. H., Lim., G., and Chang, Y. (2006). Performance analysis of scheduling algorithms for VoIP services in IEEE 802.16e systems. *IEEE 63rd Vehicular Technology Conference, 2006*, 1231 – 1235.
- Lee, H., Lee, S., and Cho, D. (2003). Mobility management based on the integration of mobile IP and session initiation protocol in next generation mobile data networks. In *IEEE 58th Vehicular Technology Conference, 2003. VTC 2003-Fall*, Volume 3, 2058–2062.
- Lee, K. and Kim, S. (2002). Optimization for adaptive bandwidth reservation in wireless multimedia networks. *Computer networks, 2002*, 38(5), 631–643.
- Li, B., Hamdi, M., Iang, D., Cao, X., and Hou, Y. (2000). QoS enabled voice support in the next generation Internet: issues, existing approaches and challenges. *IEEE Communications Magazine, 2000*, 38(4), 54–61.
- Li, B., Qin, Y., Low, C., and Gwee, C. (2007a). A survey on mobile wimax [wireless broadband access]. *IEEE Communications Magazine, 2007*, 45(12), 70–75.
- Li, B., Qin, Y., Low, C., and Gwee, C. (2007b). A survey on mobile wimax [wireless broadband access]. *IEEE Communications Magazine, 45(12)*, 70–75.
- Liu, B., Martins, P., Samhat, A., and Bertin, P. (2008). A layer 2 scheme for Inter-RAT handover between UMTS and WiMAX in tight coupling architecture. *IEEE 19th International Symposium Personal, Indoor and Mobile Radio Communications, 2008*, 1–5.
- Lopes, J. (2008). *Performance Analysis of UMTS/HSDPA/HSUPA at the Cellular Level*. Ph. D. thesis, M. Sc. Thesis, IST-UTL, Lisbon, Portugal.
- Lunden, P., Aijanen, J., Aho, K., and Ristaniemi, T. (2008). Performance of VoIP over HSDPA in mobility scenarios. *Proceedings of the 67th IEEE Vehicular Technology Conference, 2008*, 2046 – 2050.

- Motorola (2005). Space Seamless Mobility A Continuity of Experiences Across Domains, Devices and Networks.
- Nam, M., Choi, N., Seok, Y., and Choi, Y. (2004). WISE: energy-efficient interface selection on vertical handoff between 3G networks and WLANs. In *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2004*, Volume 1, 692–698.
- Nguyen-Vuong, Q., Agoulmine, N., and Ghamri-Doudane, Y. (2007). Terminal-controlled mobility management in heterogeneous wireless networks. *IEEE Communications Magazine*, 2007, 45(4), 122–129.
- Nguyen-Vuong, Q., Fiat, L., and Agoulmine, N. (2006). An architecture for UMTS-WIMAX interworking. *The 1st International Workshop on Broadband Convergence Networks, 2006. BcN 2006.*, 1–10.
- Oliveira, C., Kim, J., and Suda, T. (1998). An adaptive bandwidth reservation scheme for high-speed multimedia wireless networks. *IEEE Journal on Selected Areas in Communications*, 1998, 16(6), 858–874.
- Passito, A., Mota, E., Aguiar, R., Carvalho, L., Moura, E., Briglia, A., and Bids, I. (2005). Using an E-model implementation to evaluate speech quality in voice over 802.11 b networks with VPN/IPSec. *IEEE International Conference on Wireless Communications, Networking and Mobile Computing, 2005*, 2, 1123–1127.
- Prabhakar, B., Uysal Biyikoglu, E., and El Gamal, A. (2001). Energy-efficient transmission over a wireless link via lazy packet scheduling. In *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, Volume 1, 386–394.
- Psimogiannos, N., Sgora, A., and Vergados, D. (2010). An IMS-based network architecture for WiMAX-UMTS and WiMAX-WLAN interworking. *Computer Communications*.
- QUALCOMM (April, 2008). 3G Case Studies .
- Sarddar, D., Maity, S., Raha, A., Jana, R., Biswas, U., and Naskar, M. A RSS Based Adaptive Hand-Off Management Scheme In Heterogeneous Networks. *International Journal of Computer Science*, 2010, 7.

- Scalabrino, N., De Pellegrini, F., Chlamtac, I., Ghittino, A., and Pera, S. (2006). Performance evaluation of a WiMAX testbed under VoIP traffic. *WiN-TECH '06: Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization*, 97–98.
- Sengupta, S., Chatterjee, M., and Ganguly, S. (2008). Improving quality of VoIP streams over WiMax. *IEEE Transactions on Computers*, 2008, 57(2), 145–156.
- Sharma, G. and Purohit, G. (2011). Queue Modeling of Handoff Calls with Sub Rating in Wireless Cellular Networks. *Applied Mathematical Sciences*, 2011, 5(21), 1017–1024.
- Shinder, D. (November , 2007). <http://www.fring.com/>.
- Shuaibu, D., Syed-Yusof, S., and Fisal, N. (2007). Partition-Based Bandwidth Management for Mobile WiMAX IEEE802. 16e. *International Review on Computers and Software*, 2007, 5.
- Soldani, D. (2005). *QoS management in UMTS terrestrial radio access FDD networks*. Helsinki University of Technology.
- Sophia, Antipolis, C. (1999). Speech Processing, Transmission and Quality Aspects (STQ), Overall Transmission Plan Aspects for Telephony in a Private Network, Sophia Antipolis Cedex 0 France, ETSI.
- Sun, C., Stevens-Navarro, E., and Wong, V. (2008). A constrained MDP-based vertical handoff decision algorithm for 4G wireless networks. In *IEEE International Conference on Communications, 2008. ICC'08*, 2169–2174.
- Taniuchi, K., Ohba, Y., Fajardo, V., Das, S., Tauil, M., Cheng, Y.-H., Dutta, A., Baker, D., Yajnik, M., and Famolari, D. (2009). IEEE 802.21: Media independent handover: Features, applicability, and realization. *IEEE Communication Magazine*, 2009, 112– 120.
- Velte, A. and Velte, T. (2006). *Cisco: A Beginner's Guide* (Fourth ed.).
- Verdu, S. (2002). Spectral efficiency in the wideband regime. *IEEE Transactions Information Theory*, 48(6), 1319–1343.

- Wang, H. and Prasad, D. (2005). *End-2-End QoS Provisioning in UMTS networks*. Project Report, MSc. Mobile Communications, CPK. Aalborg University.
- Wang, W., Wang, X., and Nilsson, A. (2006). Energy-efficient bandwidth allocation in wireless networks: algorithms, analysis, and simulations. *IEEE Transactions on Wireless Communications*, 2006, 5(5), 1103–1114.
- Xu, F., Zhang, L., and Zhou, Z. (2007a). Interworking of Wimax and 3GPP networks based on IMS. *IEEE Communications Magazine*, 2007, 45(3), 144–150.
- Xu, F., Zhang, L., and Zhou, Z. (2007b). Interworking of Wimax and 3GPP networks based on IMS [IP Multimedia Systems (IMS) Infrastructure and Services]. *IEEE Communications Magazine*, 2007, 45(3), 144–150.
- Yan, X., Ahmet Sekercioglu, Y., and Narayanan, S. (2010). A survey of vertical handover decision algorithms in Fourth Generation heterogeneous wireless networks. *Computer Networks*, 2010, 54(11), 1848–1863.
- Yang, K., Gondal, I., Qiu, B., and Dooley, L. Combined SINR based vertical handoff algorithm for next generation heterogeneous wireless networks. In *IEEE Global Telecommunications Conference, GLOBECOM'07*, 4483–4487.
- Yang, S. (2007). Dynamic power saving mechanism for 3G UMTS system. *Mobile Networks and Applications*, 12(1), 5–14.
- Yang, S., Wu, J., and ROC, Y. (2007). Handoff management schemes across hybrid WiMAX and WiFi networks. In *IEEE TENCON, 2007*.
- Yang, W., Wang, Y., Tseng, Y., and Lin, B. (2009). An energy-efficient handover scheme with geographic mobility awareness in WiMAX-WiFi integrated networks. In *IEEE Wireless Communications and Networking Conference, WCNC 2009*, 1–6.
- Zeng, Q. and Agrawal, D. (2002). Handoff in wireless mobile networks. *Handbook of wireless networks and mobile computing*, 2002, 1–25.
- Dong, Qunfeng and Banerjee, Suman and Adler, Micah and Misra, Archan (2005). Minimum energy reliable paths using unreliable wireless links. *Pro-*

ceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing, 2005, 449–459.

Karl, Holger and Willig, Andreas (2007). Protocols and architectures for wireless sensor networks. *Wiley, 2007*

Munir, Arslan and Gordon-Ross, Ann (2009). An MDP-based application oriented optimal policy for wireless sensor networks. *Proceedings of the 7th IEEE/ACM international conference on Hardware/software codesign and system synthesis*, 183–192