Management framework for multiple-interface cooperation in multihomed terminal

Andi Widyo Cahyono

2010

Andi, W. C. (2010). Management framework for multiple-interface cooperation in multihomed terminal. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/21196>

<https://doi.org/10.32657/10356/21196>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**MANAGEMENT FRAMEWORK FOR MULTIPLE-
INTERFACE COOPERATION IN MULTIHOMED
TERMINAL**

ANDI WIDYO CAHYONO

SCHOOL OF COMPUTER ENGINEERING

2010

MANAGEMENT FRAMEWORK FOR MULTIPLE-
INTERFACE COOPERATION IN MULTIHOMED TERMINAL

ANDI WIDYO CAHYONO

2010

MANAGEMENT FRAMEWORK FOR MULTIPLE- INTERFACE COOPERATION IN MULTIHOMED TERMINAL

ANDI WIDYO CAHYONO

ANDI WIDYO CAHYONO

School of Computer Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Master of Engineering

2010

Abstract

Next generation network architecture features heterogeneous wireless access networks in which diverse access technologies, such as Wi-Fi, WiMAX, Bluetooth, UMTS, are deployed to offer ubiquitous services for the users. These access technologies differ in terms of coverage, frequency, capacity, and QoS support mechanisms. Given that a user's device is a multihomed one, preserving connectivity to any of the networks has been a challenge in the situation where user roams from one access network to another. This is further complicated by the existence of ongoing applications which vary in their QoS requirements. During the life time of application session, handover may be required if the user perceives any degradation in the prevailing connection and finds better support using another network interface. Regardless of whether any QoS support exists and is deployed within the network, the smart device is still required in order to optimize the user's connection.

In this thesis, we propose the Multi-homed Mobility Management Framework (MHMM) to enable multiple interface cooperation in a multihomed device. MHMM, built upon the Media Independent Handover protocol, features mobility management using Mobile Session Initiation Protocol as the mobility protocol. A virtual device agent is introduced within the framework which performs tasks relating to session management, as well as execution of policy. The results show that the framework ensures seamless communication during handover by utilizing link layer event upon which a make-before-break is implemented. In addition, both communicating parties do not experience any disruption due to IP address transparency provided by the

agent. An adaptive load balancing policy is implemented for delay sensitive applications.

We further extend the operation of MHMM by taking into account more awareness of the prevailing access networks conditions. A set of representative link layer parameters, extracted from user device's Medium Access Control (MACs), is used to classify the current access network's into more crisp states; uncongested, moderately congested, and highly congested. These states are also closely related to the throughput offered by the access networks. The device is equipped with prior knowledge which helps it to estimate the prevailing network condition and identify the network state. Load balancing schemes are then applied adaptively across the different interfaces. The new framework is called Adaptive Multiple-Interface Management Framework (AMIMF). Among the several load balancing schemes implemented in AMIMF are Round Robin, Adaptive Weighted Round Robin, Adaptive Distribution, and Hybrid scheme. Hybrid scheme results in the best performance in terms of throughput, end-to-end delay, inter-packet delay variation, and packet drop. Results show that exploiting terminal's awareness could greatly improve the user's perceived quality.

We also propose the coupling of the AMIMF with QoS support of underlying access technologies. QoS consistency is maintained accross the wireless domains as roaming occurs. Results show that the QoS-integrated framework provides efficient switching policies for effective multi-interface access as well as the ability to cater to user preferences and diverse applications characterized by their QoS requirements. Moreover, the framework's operation is mostly done at the terminal side of the communication path which requires minimum support from network entity.

Acknowledgements

First of all, I would like to express my sincere gratitude towards Associate Professor Yeo Chai Kiat for her constant guidance, support and encouragement during the course of my research study.

I would also like to thank Associate Professor Francis Lee Bu Sung for his precious insight and support during my research.

I would like to acknowledge AUN/SEED-Net for providing financial support during the course of my study.

My gratitude extends to Mr. Xia Yang and Mr. Ernest Kurniawan for their valuable ideas, support and engagement in our fruitful discussions.

I would also like to thank the staff, Ms. Chua Poo Hua and Ms. Cindy Teo Cheng Kee, and also all my friends at Centre for Multimedia and Network Technology (CeMNET) for their friendship and support. I cherish the opportunity of being in such supportive working environment. Special thanks to my colleagues, Mr. Zhong Feng and Mr. Abhik Banerjee for their encouragement and support in our discussions.

Last but not least, I would like to thank my family for their unrelenting love, prayer, and support.

Contents

Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Tables	viii
List of Figures	ix
Acronyms	xi
1 Introduction	1
1.1 Motivation and Objective of Thesis	1
1.2 Contribution of Thesis	3
1.3 Organization of Thesis	4
2 Literature Review	5
2.1 Related Work in Mobility Scheme	5
2.2 Related Work in Multiple Interfaces Usage	8
2.3 Related Work in Terminal Qos Support	10
3 MIH Based Multi-homed SIP Mobility Management (MHMM)	11
3.1 Framework for Multiple-Interface Cooperation	12

3.1.1	VDA	12
3.1.2	MIH entity	13
3.1.3	Profile entity	14
3.2	MIH IEEE 802.21	14
3.2.1	Interworking architecture	19
3.3	Multi-homed Mobile SIP	20
3.3.1	Mobility mechanism	24
3.4	Packet Distribution Policy	27
3.4.1	Background	27
3.4.2	Adaptive load balancing policy	29
3.5	Summary	33
4	Adaptive Multiple-Interface Management Framework for Multihomed Terminal (AMIMF)	34
4.1	Network Interface Selection	35
4.2	Cross Layer Information	38
4.2.1	L2 Delay	39
4.2.2	Channel Busy Time	40
4.2.3	Bandwidth Request (BW-REQ) Loss Ratio	43
4.3	Access Network Performance State	45
4.4	Packet Distribution Policy	47
4.4.1	Round Robin (RR)	47
4.4.2	Adaptive Distribution	48
4.4.3	Adaptive Weighted Round Robin (Adaptive WRR)	48
4.4.4	Hybrid Scheme	49
4.5	Simulation and Results	49
4.6	Summary	52

5	QoS Support in Multihomed Terminal	55
5.1	QoS in WLAN	56
5.2	QoS in WMAN	58
5.3	Inter-working Mechanism in QoS-enabled Terminal	59
5.3.1	WLAN $\rightarrow$ WMAN handover	62
5.3.2	WMAN $\rightarrow$ WLAN handover	64
5.4	Simulation Results	66
5.5	Summary	68
6	Conclusion	72
6.1	Conclusion	72
6.2	Recommendation for Future Work	75
	Author's Publications	77
	Bibliography	78

List of Tables

4.1	Profile on access network performance	45
4.2	Estimated load and available throughput in our framework	46
4.3	Throughput value used for adaptive WRR	48
5.1	IEEE 802.16 service class	58
5.2	Mapping of QoS for different technologies	64
5.3	AC setting in terminal	67

List of Figures

3.1	Framework for multiple wireless interfaces management	13
3.2	Media independent handover framework	15
3.3	MIH signaling flow diagram: initial stage	17
3.4	Overview of multi-homed mobile SIP	21
3.5	SIP signaling: registration	22
3.6	SIP signaling: session establishment	23
3.7	Terminal moves out the coverage of WLAN access network	25
3.8	Illustration of access network coverage power threshold	25
3.9	Throughput when handover occurs	27
3.10	Illustration of path selection	29
3.11	Policy impact on application's throughput	31
3.12	Policy impact on application's end to end delay	32
3.13	Policy impact on application's inter-packet delay variation	33
4.1	Architecture of multihomed terminal equipped with three interfaces: Wi-Fi, WiMAX, and UMTS	36
4.2	Normalized load vs L2 delay in Wi-Fi	40
4.3	Normalized load vs L2 delay in WiMAX	41
4.4	Normalized load vs channel busy time in Wi-Fi	42
4.5	Normalized load vs BW-REQ loss ratio in WiMAX	44
4.6	Policy impact on (a) throughput ; (b) number of packets received . .	53
4.7	Policy impact on (a) end-to-end delay ; (b) inter-packet delay variation	54

5.1	Access Categories in EDCA mode	56
5.2	Illustration of WLAN and WMAN network integration	59
5.3	Policy I: Delay in dual-interface terminal	69
5.4	Policy I: Throughput of network interfaces in dual-interface terminal	69
5.5	Policy I: Throughput of applications in dual-interface terminal	70
5.6	Policy II: Delay in dual-interface terminal	70
5.7	Policy II: Throughput of network interfaces in dual-interface terminal	71
5.8	Policy II: Throughput of applications in dual-interface terminal . . .	71

Acronyms

AC Access Category.

AMIMF Adaptive Multiple-Interface Management Framework

AP Access Point

BE Best Effort

BS Base Station

BW-REQ Bandwidth Request

CBR Constant Base Rate

CN Correspondent Node

CoA Care-of-Address

HA Home Agent

HoA Home Address

L2 Link Layer

MAC Medium Access Control

MADM Multiple Attribute Decision Making

MHMM MIH Based Multi-homed SIP Mobility Management

MICS Media Independent Command Service

MIES Media Independent Event Service

MIH Media Independent Handover

MIHF Media Independent Handover Function

MIIS Media Independent Information Service

MIP Mobile IP

MIPv6 Mobile Ipv6

MN Mobile Node

MSIP Mobile SIP

NGN Next Generation Network

nrtPS Non-Real-Time Polling Service

PMP point-to-multipoint

PoA Point of Attachment

PoS Point of Service

QoS Quality of Service

RDA Reference Destination Address

RR Round Robin

RSA Reference Source Address

RSS Received Signal Strength

rtPS Real-Time Polling Service

SAP Service Access Point

SAW Simple Additive Weighting

SIP Session Initiation Protocol

SS Subscriber Station

UA User Agent

UGS Unsolicited Grant Service

URI Uniform Resource Identifier

VDA Virtual Device Agent

WMAN Wireless Metropolitan Area Network

WRR Weighted Round Robin

Chapter 1

Introduction

Next generation wireless network has been envisioned as having diverse access network technologies, such as Wireless Local Area Network (WLAN), Wireless Metropolitan Area Network (WMAN), 3G UMTS, integrated to provide ubiquitous services to users. On the other hand, proliferation of devices equipped with several network interfaces has driven a demand for user's connectivity for all types of applications. Mass production of user devices such as laptop, PDA, Smart phone has lowered the cost of providing multihoming¹ feature for end users. Thus, the existence of many service providers and various access technologies, which complement one another in terms of offered network capacity and coverage, give the users flexibility to choose their own service providers and yet at the same time, benefit from multihoming capability according to their preference or policy.

1.1 Motivation and Objective of Thesis

WLAN, based on IEEE 802.11 a/b/g standard has been deployed widely and offers high bandwidth connection over small coverage area for single access point. Meanwhile, WMAN WiMAX, based on IEEE 802.16 d/e standard, provides high capacity

¹multihomed refers to having multiple network interfaces [1]

CHAPTER 1. INTRODUCTION

last-mile connectivity whereas UMTS offers lower bandwidth over a wide geographical area. Communication infrastructure that integrates these technologies poses several challenges that should be carefully addressed at the end user. Thus, the framework deployed within the device should take into consideration the following issues.

First issue concerns with a case of user loosing connectivity of one access technology or moving into the coverage of other network. Thus, the framework needs to feature mobility management. Mobility management comprises handover management and location management [2]. The objective of mobility management is to ensure seamless service continuity for the user. This also deals with service optimization in which ongoing session is moved from one network interface to another due to performance degradation of prevailing access network. The situation is likely to occur owing to the fluctuating characteristic of wireless link. Location management deals with enabling the network entity to locate the current attachment point of the user. It comprises location registration and location update.

Second issue deals with how the user optimizes the usage of his network interfaces. Having multihomed terminal on the end user raises a challenge on how to fully exploit terminal's capability that users may benefit from features such as path redundancy, load balancing, and available aggregate bandwidth.

In this thesis, we aim to tackle the abovementioned issues by proposing a framework which resides at the user device. We put most of the intelligence into the terminal rather than into the network for the reason that user has more flexibility and control to choose connectivity that maximizes his subjective perceived quality. Hence, the framework requires no modification to the network side as it relies on already established component or function in the network infrastructure, allowing the deployment of various access networks to be made independently. The final objective of the framework would be enhancement of user's experience when using the network service by offering better Quality of Service (QoS) support.

CHAPTER 1. INTRODUCTION

1.2 Contribution of Thesis

In this thesis, we propose a terminal-side multiple interfaces management framework that incorporates a scheme for mobility and session management as well as to enable cooperation of multiple wireless interfaces to support user's applications. The framework is QoS-aware by taking into account the QoS parameters both from the underlying network and the application. The framework supports vertical handover as well as horizontal handover to enable user to seamlessly redirect its communication session in the midst of the session.

The main contributions of this thesis are as follows:

1. Development of MIH-based Multi-homed Session Initiation Protocol (SIP) Mobility Management framework (MHMM) in a dual-interface (Wi-Fi - WiMAX) user device. Through MHMM, long handover delay incurred by SIP as application layer protocol can be anticipated by exploiting MIH's Media Independent Event Service. Hence, MHMM supports seamless communication experience for the user. MHMM also features Multi-homed SIP which enables registration of multiple addresses associated with a user agent. In addition, MHMM incorporates adaptive load balancing policy for optimizing the multiple interfaces usage especially for delay sensitive application.
2. Development of Adaptive Multiple-Interface Management framework (AMIMF) in a triple-interface (Wi-Fi - WiMAX - UMTS) user device. AMIMF equips the device with a prior set of knowledge which enables the device to intelligently estimate prevailing access network's condition and infer the network state. AMIMF implements load balancing schemes, such as Round Robin, Adaptive Weighted Round Robin, Adaptive Distribution, and Hybrid scheme, to govern the outgoing transmission of packets. Simulation results show that Hybrid scheme yields the best performance. AMIMF is beneficial especially for application characterized by long life time.

CHAPTER 1. INTRODUCTION

3. Integration of modified AMIMF with QoS support from user device's underlying access technologies. The QoS-integrated AMIMF, implemented in dual-interface (Wi-Fi - WiMAX) device, features the following innovation: accurate mapping of application QoS to the QoS mechanisms of the different access technologies and between technologies and efficient interface switching policies. Hence, the framework is capable of entertaining diverse applications according to their QoS requirements as well as catering to different user preferences.

1.3 Organization of Thesis

The remaining of the thesis is organized as follows:

- Chapter 2 presents several related works on mobility scheme, multiple interfaces usage, and terminal QoS support.
- Chapter 3 presents the MIH Based Multi-homed SIP Mobility Management framework, namely MHMM.
- Chapter 4 presents the Adaptive Multiple-Interface Management Framework, namely AMIMF, which is an extension of MHMM.
- Chapter 5 presents the integration of AMIMF with link layer QoS mechanism.
- Chapter 6 concludes the thesis and provides a recommendation for future work.

Chapter 2

Literature Review

2.1 Related Work in Mobility Scheme

Li et al. [3] have proposed a multi-interface scheme to support handover. The scheme adopts the implementation of IEEE 802.21 Media Independent Handover (MIH) [4] to aid network interface selection. The authors simulate a dual-interface mobile host, all are of IEEE 802.11 technology. Each network interface can work in either primary mode or standby mode. However, only the interface in primary mode is activated for carrying traffic streams as well as signaling. Another interface is in standby mode, which is only allowed to receive signaling from a new point of attachment (PoA). In the current scheme, the decision to perform handover is solely based on signal strength parameter.

You et al. [5], aimed at supporting reliable accessibility and vertical handover, have proposed an extension of multiple home address registration mechanism in mobile IPv6 (MIPv6). The idea is a mobile node equipped with multiple interfaces can have multiple care-of addresses (CoAs) and multiple home addresses (HoAs) and needs to register all these addresses. The optimization of the registration mechanism is achieved by modifying the binding update and adding flag with mobility sub option to provide multiple CoAs registration. The authors propose extended multiple home

CHAPTER 2. LITERATURE REVIEW

address registration mechanism which handles multiple home address like care-of address. However, in the case of vertical handover, large overhead is incurred as the mechanism needs two different bi-directional tunneling (for example, between second Home Agent (HA2) and HA1, and between HA1 and MN). The work only provides the concept of using multiple interfaces simultaneously to support reliability and does not consider using them for the purpose of load balancing whenever applicable.

Jover et al. [6] propose a Session Initiation Protocol (SIP) mid-call mobility scheme within the same domain. This scheme works without SIP re-invitation, but instead only requires the client to register its new IP address for the session redirection. This is beneficial in the case where a client is involved in more than one session, thus preventing large signaling overhead between corresponding clients. This scheme also gives network more control over the traffic. However, the scheme is suitable only for SIP operation in proxy mode, hence increasing the packet delay and possibility of server bottleneck since all packets need to be relayed through the SIP Proxy server.

A new architecture of multiple interface IPv6 end-user terminal has been presented in [7]. The proposed architecture supports network access optimization for end users roaming from one access network to another of different technology. It relies on tight interactions amongst upper and lower layers and provides a platform to support adaptive application due to the fluctuation of network resources.

Fikouras et al. [8] have proposed a protocol extension of Mobile IPv4, called Filters for Mobile IP. The extension enables mobile nodes to associate one or more set of rules (filters) with mobility bindings during registration with the HA. The filtering information is attached to the registration request. It specifies the type of filtering criteria and the format of control data that should be communicated between Mobile IP entities. Based on the filter activated, matching flows can be processed accordingly. By this scheme, it is possible to have only one HoA for all interfaces and direct the flows to different CoAs. The solution also enables the localized management of flow mobility and prevents signaling between end to end communication peers. However,

CHAPTER 2. LITERATURE REVIEW

the solution only specifies how to distribute incoming flows, i.e. when the flows go from correspondent node (CN) to mobile node through HA, and not vice versa.

Lampropoulos et al. [9] have identified the potential of using MIH standard for service provision and for minimizing disruption during an inter-technology handover. The authors point that the translation of resource information across different technology networks is still an open issue in the standard. A case study on vertical handover between WiMAX and GPRS network has been presented in the work. The scenario shows a network-controlled handover in which the WiMAX point of service (PoS) initiates the signaling to Media Independent Handover Function (MIHF) user entity and GPRS PoS during initiation, preparation, and execution phase of handover.

Eastwood et al. [10] have also shown the importance of MIH to support seamless mobility between an 802.11VHT and an 802.16m network which, combined together, is believed to be likely proposed for IMT-Advanced 4G. MIH provides mechanism to prepare the target networks before handover execution phase occurs, hence reducing latency. The handover flow depicted in the case study is of terminal-controlled type in which mobility decision is made by an MIH-enabled mobile node. Based on the scenario, the mobile node (MN), which is initially connected to the IEEE 802.16m access network and moving at a high speed, should be intelligent enough to detect its decreasing velocity and prepare for a possible handover to an IEEE 802.11VHT access network to obtain higher throughput. MN also evaluates the suitability of each candidate access network to support the current session and decides which candidate will be the target network.

An experimental work on handover performance in terminal supporting dual Wi-Fi interface has been conducted in [11]. The result shows that MIH-assisted handover greatly improves packet loss and delay time compared to non-MIH-assisted handover. SIP is used for the mobility protocol, and dynamic IP binding is achieved through MN sending SIP Re-Invite message to the SIP proxy server whenever it is assigned a new IP address. However, the scheme does not operate transparently and application

CHAPTER 2. LITERATURE REVIEW

layer is still aware of the IP address change.

A terminal-based IP mobility support scheme that enables personal and terminal mobility for IP-based application has been proposed in [12] [13]. A user agent is introduced to manage changes in IP address and an IP-to-IP address mapping scheme is performed at network layer to ensure transparency for upper layers. The scheme does not need a permanent IP address nor a home server as well as tunneling mechanism. The scheme supports handover management and location management is provided through a SIP server functioning both as registrar and redirect server.

Our proposed framework is based on the mobility scheme in [12] [13]. We have however, extended the said mobility scheme in our proposed framework to support multihoming in the user device, where it is tailored specially to govern the outgoing transmission of packets.

2.2 Related Work in Multiple Interfaces Usage

Endeavours for multiple interfaces usage have appeared in several literatures at various layers of the protocol stacks. At the transport layer, LS-SCTP [14], which is an extension of SCTP, aggregates multiple paths for the provision of bandwidth intensive real-time application. R-MTP [15], being a complement of TCP, constitutes a rate-based transport protocol for multiplexing a single application data stream across multiple network interfaces. R-MTP can differentiate the cause of packet loss due to either congestion or transmission errors. However the scheme assumes that the network provides a stable throughput. Parallel TCP (pTCP) [16] is another end-to-end transport layer protocol that supports stripping of connections. pTCP constitutes a wrapper for Stripping manager (SM) and TCP-virtual (TCP-v). SM handles aggregate connection functionalities and each TCP-v handles each micro-flow of a stripped connection. Even though pTCP enables bandwidth aggregation through multiple interfaces, it introduces overhead as it requires multiple connection establishment in

CHAPTER 2. LITERATURE REVIEW

the case of stripping application into multiple micro-flows.

On the network layer, several solutions are proposed in [17], [18], [19]. Lin et al. [17] propose a dynamic bandwidth aggregation (DBA) approach to aggregate the bandwidths and to share loads between multiple wireless interfaces. An entity called DBA proxy is located at the network to split the connection between MN-CN and offers bandwidth aggregation service for both inbound and outbound traffic. However, the scheme does not consider congestion in network and simulates the dynamics of wireless link by varying the throughput of the link. Huang et al. [18] propose a system that supports multi-homed communication. A load-sharing scheme is deployed based on Weighted Round Robin link selection and Jump-ahead packet scheduling. This scheme also assumes static bandwidth for each network link. A framework called INTELiCON [19] operates transparently to both the application and the link layer. The framework adopts adaptive transmission strategies over multiple interfaces. The selection and evaluation of the strategy depends on the current network state which is communicated by piggy-backing some metrics (e.g. packet loss, delay) into the IP packet.

Solution at the link layer, such as in [20] is suitable for diverse links belonging to the same access technology. As multihomed user device usually comprises interfaces of different technologies (eg. IEEE 802.11 and Ethernet or UMTS), the scheme does not suit well in the context of heterogeneous wireless networks.

Several works on MIH and its enhancement can be found in literatures, such as [11], [21], [22], [23]. However the works are mainly for optimization of vertical handover across heterogeneous networks. Hence those works do not exploit the multiple interfaces to enhance user's perceived quality.

Our framework exploits MIH and provides terminal's awareness of its prevailing access network condition. User is given flexibility to switch his connection over the terminal's interfaces should the device experiences performance degradation in the network.

CHAPTER 2. LITERATURE REVIEW

2.3 Related Work in Terminal QoS Support

A design to integrate WLAN and WMAN link layer QoS into next generation network (NGN) is proposed in [24]. A mapping between the WLAN and WMAN link layer QoS classes and NGN DiffServ class is presented. The work is built upon the Moby Dick architecture [25] which provides QoS at the network level, and further extends it with the signaling required to set up QoS-enabled connection at the link layer. However, the work does not consider QoS handover management, i.e. how the session is maintained with equivalent service class when the user roams accross different access technologies. Ali-Yahiya et al. [26] propose WLAN/WMAN integration architecture with focus on link layer QoS provisioning support for user's ongoing sessions when moving from WLAN to WMAN. The work investigates related issues such as QoS mapping and vertical handover signaling. As the work assumes that all sessions originate from the WLAN cell, network interface selection during session initiation is never within its scope. QoS framework for heterogeneous network which includes WLAN/WMAN service class mapping, call admission control (CAC) and scheduling in WMAN base station is proposed in [27]. However the work also does not provide a discussion on network interface selection and decision criteria when the user device performs vertical handover.

We propose integration of our proposed framework with link layer (L2) QoS mechanism, of both WLAN and WMAN, to enhance the performance of various applications over the wireless links. We also present scenario of handover both from and to WLAN/WMAN which utilizes L2 metrics obtained from MIH module as the decision criteria.

Chapter 3

MIH Based Multi-homed SIP Mobility Management (MHMM)

Seamless handover has been an indispensable requirement for NGN. To ensure the seamlessness, low handover latency and same level of service after handover constitute the key elements [9]. There are existing two dominant solutions for mobility management upon which many works have been based. At the network layer, Mobile IP (MIP) [28] protocol supports mobility across different IP domains. MIP provides transparent mobility that any changes of terminal's IP address is hidden from application layer. However, using MIP imposes several drawbacks, including triangle routing, overhead by IP-in-IP header encapsulation, and the need for permanent HoA. MIP for IPv6 (MIPv6) [29] has emerged to overcome several limitations of MIP. MIPv6 does not incur triangle routing anymore by a binding update mechanism so that CN builds a direct path to source node. However, since the binding update from source node is still tunneled through source node's HA, the handover delay is relatively high and hence incurs significant packet loss. Another protocol, i.e. SIP [30], has emerged as a viable solution for IP mobility. Initially proposed by Internet Engineering Task Force (IETF) as a signaling protocol for multimedia session setup, SIP supports personal and session mobility as well. However, since the protocol lies in the application

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

layer, SIP intuitively suffers from longer handover latency compared to MIP [31].

This chapter proposes a MIH Based Multi-homed SIP Mobility Management (MHMM) framework which features mobility management and cooperation of multiple wireless interfaces by taking into account the QoS parameters from both the network and the application. The framework also supports vertical handover to enable the user to seamlessly redirect his communication session in the midst of the session. Since the terminal has more than one network interface, a *make-before-break* mechanism or commonly known as soft handover, is deployed to prevent packet loss during handover period.

3.1 Framework for Multiple-Interface Cooperation

Our proposed framework for multiple wireless interfaces management is shown in Fig. 3.1. Currently, we assume that the device is equipped with two network interfaces of different technologies, i.e. IEEE 802.11b WLAN and IEEE 802.16 WMAN. The proposed framework comprises three main components, namely Virtual Device Agent (VDA), MIH entity, and Profile entity.

3.1.1 VDA

From the multihomed terminal's point of view, VDA acts as a virtual interface. VDA provides abstraction and hides the heterogeneity of physical network interfaces from the application layer. Thus, application, being aware only of this interface, relays the packet to VDA as well as receives packets from it. VDA constitutes the intelligent part of the terminal which handles functionalities such as network interface selection and flow or packet redirection.

VDA has the following functionalities :

1. Keep a table on the mapping between packet flow and network interface

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Figure 3.1: Framework for multiple wireless interfaces management

2. Execute policies for packet flow
3. Redirect flow in case of handover

VDA implements IP-to-IP mapping mechanism [12] [13] to ensure IP address transparency above the network layer. This mechanism hides IP address change from the application, thus avoiding disruption in communication session. In order to maintain end-to-end connection, VDA is deployed both at the communicating terminal and its CN.

3.1.2 MIH entity

MIH resides between the data link and the network layers, or often referred to as layer 2.5. It provides MIH IEEE 802.21 functionalities. For MIH simulation, we utilize a

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

patch from [32].

3.1.3 Profile entity

Profile storage saves information from both the application and the network layer. Information related to the application layer may include maximum upper delay bound and minimum bandwidth requirement of application. These metrics are important to provide better support for delay sensitive applications such as audio and video conference. According to the user policy and requirement of application, VDA checks the available interfaces to determine which of them are suitable as candidates for the interface selection. The metrics can also be extended to include network usage cost. This is considerably important if users want to make preference based on the amount they would like to spend if the network service is not free. Information related to link layer includes delay and frame loss. This information is sent to the upper layer using `MIH_Link_Parameters_Report` primitive. In order to reflect the dynamic condition of access networks, link layer information in the storage is updated at a regular interval.

By nature, efficient cooperation between different access networks is hard to achieve since specific access network technology has its own mobility mechanism, QoS support, and security requirement. In the following sections, we give an overview of the technologies which are combined to offer a solution for seamless handover.

3.2 MIH IEEE 802.21

MIH is a standard developed by IEEE 802.21 working group to enable the handover between various access network technologies including both 802 (802.11, 802.16) and non 802 family (cellular 3GPP) networks. It provides a generic framework to facilitate link layer control and communicate heterogeneous network information to upper layers with regards to optimizing handover. The core functionality of MIH lies in the MIHF that provides abstraction of link layers to higher-layer entities, i.e. MIH-users (ex.

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

VDA). Media independent service access point (SAP) interfaces MIH-users with the MIHF, whereas the media dependent SAP provides interface for specific link layer technology. MIHF, located both in terminal and the network entity protocol stack, provides the following three services:

- Media-independent event service (MIES)
- Media-independent command service (MICS)
- Media-independent information service (MIIS)

Figure 3.2: Media independent handover framework

Figure 3.2 shows the media independent handover framework along with the flow direction of MIES, MICS, and MIIS. MIES provides event detection and reporting from lower layer to upper layer. Some of the events defined in IEEE 802.21 include link detection, link up, link going down, and link down indication. MIES also enables terminal to report link status measurement in periodic manner or when certain preset

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

thresholds are crossed. This is realized using link parameter report event in which dynamic link layer information such as delay and frame loss are delivered to MIH-users. These information can be used intelligently to support the decision whether the terminal performs handover. MICS provides a set of commands for MIH-users to control the link layer for the handover purposes. Example of commands include scanning of available base station (BS) or access point (AP) and configuration of network device. MIIS provides information to assist the handover process. Using this service, MIHF in mobile node is able to obtain information, both static and dynamic with regards to available neighbouring networks through a *request-response* mechanism with an information server (IS). IS resides in network entities such as BS or AP so that it can be contacted easily by all the clients in the access network domain. Static information comprises, for example, MAC address and channel information of PoA, whereas dynamic information includes link layer parameters whose value changes over time such as, data rate, and instantaneous throughput.

MIH standard clearly defines handover mechanism in initiation and preparation phase, whereas for handover execution phase, IEEE 802.21 depends on higher layer mobility management, for example MIP and Mobile SIP (MSIP). Functionalities of IEEE 802.21 enables two scenarios of handover, i.e. network-controlled handover (NCHO) and terminal-controlled handover (TCHO). In NCHO, network initiates control of handover and makes decision of network selection in coordination with terminal. In TCHO, terminal initiates control as well as makes all decisions related to handover. TCHO is more suitable in an environment where diverse access networks are completely independent and belong to non-collaborating operators. We focus our work on TCHO. For illustration of network-controlled handover between WiMAX and GPRS network, refer to [9].

In our scenario, terminal is initially located in the overlap area of both WLAN (Wi-Fi) and WMAN (WiMAX) access networks. At this point, terminal maintains all connections to both networks. The corresponding MIH signaling, as shown in

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Figure 3.3: MIH signaling flow diagram: initial stage

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Fig. 3.3, can be explained as follows:

1. MIH_Get_Status primitive is issued by MIH-users which requests MIHF to poll the lower layer to get information on the status of the link. This process returns operation mode and link identifier information of the corresponding network interface.
2. In infrastructure mode of IEEE 802.11, in order to inform hosts about its presence, an AP broadcasts beacon frame at a specific interval. With the beacon received in its Wi-Fi interface, terminal can detect whether the beacon comes from the same AP it is connected to. If yes, terminal updates its beacon timer. If not, a Link_Detected event will be triggered and sent to MIH-users, which subsequently commands MIHF to establish L2 connection with the AP through Link_Connect command. An Association *request-response* mechanism is performed for connection establishment. Using the mechanism, admission control into WLAN network can be performed. If the AP receives the Association request and decides that it can provide service to terminal, it will return an Association Response frame indicating the success of connection establishment. The success of this process will be followed by a Link_Up event indication. In point-to-multipoint (PMP) mode of IEEE 802.16 network, a Link_Detected event will be triggered when terminal's WiMAX interface receives downlink map (DL_MAP) for synchronization with BS. Then, MIH-users issue Link_Connect command to MIHF. After receiving uplink channel descriptor (UCD) parameter, terminal proceeds with the initial ranging process on its WiMAX interface and start the contention process to notify the BS of its entry into the system. Upon a successful Registration *request-response* scheme, Link_Up indication on WiMAX interface will be triggered.
3. Upon receiving Link_Up indication, MIHF in the terminal uses MIH primitive called MIH_Capability_Discover.request to discover AP's and BS's MIH

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

capability information of their link layers. This primitive contains SupportedMihEventList, SupportedMihCommandList, SupportedMihTransportList, and SupportedISQueryType to enable the AP and BS's MIHF to recognize the terminal's MIHF capability. MIHF entities in IEEE 802.11 AP and IEEE 802.16 BS will respond with MIH_Capability_Discover.confirm primitive containing the value of the same parameter. Therefore, using this primitive, peer MIHF entities can recognize each other's MIH capability using one MIH protocol message transaction.

4. MIH registration provides a basis for network controlled inter-technology handover framework. This mechanism creates a pairing between terminal MIH entity and network MIH entity, thus enabling the network MIH entity to provide any MIH services to the terminal. To perform the registration to MIHF in IEEE 802.11 AP as well as IEEE 802.16 BS, terminal's MIHF uses MIH_Register.Request primitive.
5. In addition to regular update of link layer measurement in profile database, MIH-users can request MIHF to do configuration of link parameter threshold through MIH_Configure_Threshold primitive. This can be applied if there is an application running in the terminal with strict requirement, e.g. upper delay bound, and minimum throughput. Using this primitive, upper layer can be notified of lower layer changes when the measured delay or throughput crosses the preset threshold.

3.2.1 Interworking architecture

Heterogeneity of NGN in terms of wireless access technologies and service entails the need of interworking framework to achieve convergence with IP technology being the backbone of the core network. 3GPP has presented two interworking architectures for 3G cellular and WLAN network: *tight-coupling* and *loose-coupling* [33]. In *tight-*

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

coupling architecture, WLAN appears to the 3G core network as one of its radio access network. The architecture is advantageous in terms of QoS control for time-sensitive application. However, *tight-coupling* incurs high cost and complexity as WLAN should emulate all functionality existing in cellular radio access network. This necessitates the WLAN to implement 3G's mobility management and authentication, hence requires major modification to the hardware and software of WLAN device. In addition, the architecture exhibits following disadvantages: both WLAN and 3G cellular network should be owned by the same wireless operator; bulky traffic of WLAN that flows into 3G core network may lead to capacity problem.

In the *loose-coupling* architecture, the access networks are deployed independently, allowing separate data paths between WLAN and 3G network. This interworking is advantageous in terms of lower cost deployment and less complexity compared to *tight-coupling*. It does not require both networks to be under the same operator and requires less modification to the existing device. In this thesis, we assume that the heterogeneous wireless networks are loosely-coupled where these networks are connected via the core network. However, this architecture poses a drawback. As data paths are quite distinct and separate, signaling for handover is more costly, leading to possibly higher handover latency.

MIH facilitates *loose-coupling* interworking because it does not need strong interdependency between deployed access networks. To alleviate handover latency problem, the framework utilizes L2 event notifications and controls provided by MIH to help anticipate the handover.

3.3 Multi-homed Mobile SIP

SIP is an application layer control protocol used to establish, modify, and terminate multimedia session. SIP works independently of underlying transport protocol. In the basic operation of SIP, several logical entities are defined according to the speci-

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Figure 3.4: Overview of multi-homed mobile SIP

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

fication, such as user agent (UA), redirect server, proxy server, and registrar. Session initiation and termination requests originate from the user agent which resides in a host. Redirect server receives request from user agent (caller) and returns response indicating the address of the destination (callee) to where a caller should send his request. Proxy server, either in stateful or stateless operation mode, forwards the incoming request from caller to another server on the behalf of the caller. Registrar keeps the Uniform Resource Identifier (URI) and contact address of all registered UA in its domain. In our scenario, functionalities of both redirect server and registrar reside in a single server, referred to as SIP server [12]. In SIP, each user agent is identified by a URI. This SIP URI takes a form similar to an email address that contains two parts; the user part and the host part, separated by an “@”, e.g. tim@domain.com. Using this approach, users can maintain a single visible identifier regardless of their location in the network. Hence, SIP supports personal and terminal mobility, in which the dual meaning of IP address as both end point identifier and location identifier is decoupled.

Figure 3.5: SIP signaling: registration

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Our scheme, as depicted Fig. 3.4, preserves the original SIP entities. However the scheme extends the SIP operation to include support for multihoming. Fig. 3.5 exemplifies the SIP registration process. Both UA in terminal and CN need to register with their SIP server (SIP server A and SIP server B respectively) concerning their URI and contact address. However, since the terminal has two active network interfaces at the time of registration, it sends a list of contacts in the registration message containing its URI and two CoAs through which it can be contacted. For this purpose, we set the bit in the message indicating that multiple address registration is being used. The interface which has the lower link layer delay will be placed first in the ordered list.

Figure 3.6: SIP signaling: session establishment

Prior to communicating, terminal and CN need to establish a SIP session. This

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

is illustrated in Fig. 3.6. In the case of terminal initiating the session, terminal sends invitation message to its SIP server. Based on CN's SIP URI provided by terminal, SIP server A does a lookup to see if CN is registered in its domain. If not, the server sends DNS query message to DNS server. After getting a reply, SIP server A forwards the invitation message to SIP server B which is in charge of CN's SIP URI domain. SIP server B replies to SIP server A, which subsequently forwards the reply to the terminal. Upon knowing the address of CN, terminal sends invitation message to CN together with its newly updated list of CoA. By default, CN selects the first CoA in the list to establish a session with terminal. By taking into account the dynamics of access networks, it is possible that for different SIP sessions, the terminal's contact address to be used for communication is different, thus allowing for SIP session load balancing. The pair of IP addresses used during session establishment will be maintained by VDA, which is located in both terminal and CN, as reference source and reference destination address, denoted as RSA and RDA, in the respective hosts [12]. In addition, VDA also records the port number as the application identifier.

3.3.1 Mobility mechanism

Owing to mobility, user may roam out of one access network coverage as shown in Fig. 3.7. After being in the overlap area, terminal moves away from Wi-Fi AP into WiMAX network coverage. Right before terminal reaches the border of Wi-Fi access network, terminal will detect continuous signal strength degradation on its Wi-Fi link. This condition will trigger a link going down indication. Fig. 3.8 illustrates the coverage of access network along with its power threshold. The generation of link going down event is based on the power level of the received packet. For this event generation we set coefficient of *alpha* to be greater than 1. MIH Link_Going_Down event helps terminal prepare for handover execution before the actual link disruption occurs. The sequence of events in Fig. 3.7 can be explained as follows :

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Figure 3.7: Terminal moves out the coverage of WLAN access network

Figure 3.8: Illustration of access network coverage power threshold

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

1. Before handover occurs, communication path between CN and terminal exists through WLAN (Wi-Fi) network.
2. After MIH-users in terminal is notified by MIHF of the link going down indication on the Wi-Fi link, MIH-users (i.e. VDA) checks whether the probability (the probability goes higher as terminal approaches the border of AP coverage since the received packet's power goes lower) to be disconnected from Wi-Fi AP is greater than the preset threshold. If yes, terminal will check its already established SIP session and send SIP Re-Invite message to CN via its Wi-Fi interface. Since the Wi-Fi interface is considered to be down soon, terminal attaches the updated list of its addresses that can be used, which in this case only contains the IP address of terminal's WiMAX interface. In the Re-Invite message, SIP shall preserve the same call identifier value as the one in the previous session establishment.
3. Upon receiving the Re-Invite message from the terminal, CN's SIP UA obtains the call identifier information and modifies the corresponding session using the new specified contact address. Finally, CN redirects its communication flow to use the terminal's WiMAX interface address. Note that both applications in the terminal and CN are not aware of the address changes used for the communication. Since the RSA and RDA are static during the whole session, their values in the terminal and CN do not change and the VDA in CN will update its Current Destination Address (CDA) value with the terminal's WiMAX interface address.
4. Terminal sends SIP Re-Register message to its SIP server containing the updated contact address.

Fig. 3.9 shows that employing MIH feature for MSIP helps the terminal smoothen the handover process and reduce the packet loss. This is true in the case of a mobile

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Figure 3.9: Throughput when handover occurs

node having more than one network interface where a *make-before-break* mechanism can be implemented.

3.4 Packet Distribution Policy

3.4.1 Background

Packet distribution policy is one factor that determines the level of satisfaction perceived by users when using the network service. In condition where users have access to several networks, applying policy to balance the traffic flows among available network interfaces is beneficial. One of the prominent issues in load balancing is how to prevent out-of-order packet arrival at the receiver [34].

Two common load balancing schemes exist, i.e. packet-based and flow-based load balancing. Packet-based load balancing splits traffic at the finest granularity. This, for example, can be achieved using round robin scheduling. This type of balanc-

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

ing is efficient in terms of network resource usage given that the delay introduced by each path is similar. On the contrary, applying this scheme in heterogeneous networks where delay varies among paths can cause out-of-order arrival problem at the receiver. To handle the packet reordering, one solution could be to deploy a re-sequencer at the receiver. This is costly and inefficient if a high degree of re-ordering occurs since it is likely to violate upper delay bound, which is not good for delay sensitive application like VoIP. For TCP-based application, even though it is immune to out-of-order problem, high degree of re-ordering could lead to severe performance degradation since TCP may interpret out-of-order arrival as congestion indication in network and therefore will initiate slow start mechanism to avoid congestion collapse.

On the other hand, flow-based load balancing assigns packet based on the flow, which could be identified by the tuple of source-destination IP address and source-destination port number. Once the network interface to be used is decided, the mapping between traffic flow and network interface is saved for future reference. In our framework, we follow the approach for flow-based load balancing. VDA, when first receives packet from upper layer, extracts flow identifier of the packet and looks up the flow mapping table to see if there is already an interface associated with the flow. If not, VDA initiates interface selection decision algorithm. The interface list resulting from the interface selection will be mapped with the flow identifier in the flow mapping table. To maintain connection state in case of flow redirection using another interface, the first ordered interface in the list is used as the static address which remains during the whole session. Since all packets belonging to a certain flow are sent to a fixed path in flow-based load balancing scheme, packet reordering is not an issue. However, in the case involving numerous traffic flows in networks with different rates and duration characteristic, applying flow-based load balancing scheme is disadvantageous as it could lead to improper balance among paths and even throughput drop if certain network is heavily congested. Since we aim to take advantage of joint resource of different radio access networks as well as to reduce

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

out-of-order packet to a minimum, we define another policy which is named adaptive load balancing policy.

3.4.2 Adaptive load balancing policy

Figure 3.10: Illustration of path selection

Our scheme adopts adaptive load balancing policy to split up the traffic while keeping the out-of-order packet to a minimum. This policy is not unique in the sense that it is applicable to all kinds of application. The policy is applicable especially for traffic characterized with burst and idle period (or On and Off period). One example would be audio conference application. The idea behind this policy is simple and can be illustrated in Fig. 3.10. When burst 1 arrives at router 1, router 1 may forward the burst to an arbitrary path, be it path 1 or path 2 since it is the first burst of the flow. In the example, router 1 selects path 1 to send out the burst. Burst 1 is then followed by an idle period. If the burst inter-arrival time is greater than the delta delay between path 1 and path 2 (in the example, the value is 20 ms), router 1 may also forward the subsequent burst arbitrarily to any path. Since at that time burst 1 already reaches the destination router, out-of-order arrival problem is not an issue at the destination router.

We apply similar approach towards the adaptive load balancing policy. The traffic splitting is done by cutting off the flow when the packet inter-arrival time (or in our scenario, inter-generation time) is bigger than the threshold parameter ϕ , which is

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

defined as the maximum delay difference among available paths. In our scenario, $\wp$ is defined as the link layer delay difference between IEEE 802.11 and 802.16 access network. The mechanism is formulated as (3.1).

$$Action = \begin{cases} forward\ to\ previous\ path & \Delta t < \wp \\ execute\ path\ selection & \Delta t \geq \wp \end{cases} \quad (3.1)$$

where,

$$\Delta t = t_i - t_{i-1}$$

t_i : arrival time of packet i^{th} at VDA

t_{i-1} : arrival time of packet $(i - 1)^{th}$ at VDA

When packet's inter-arrival time is smaller than the threshold, the packet must be forwarded to the same path as the previous packet. Otherwise, the packet is forwarded to the path resulting from the interface selection algorithm. Since our goals are to minimize out-of-order packet problem as well as to enhance user's experience (e.g. reducing end to end delay), forwarding the burst arbitrarily is not enough. Therefore we compute merit value of each access network to help VDA in the interface selection process. We rely on MIH to obtain link layer metric, i.e. delay and frame loss, from all active interfaces. Since there is more than one metric type, merit value of access network is computed using Multiple Attribute Decision Making (MADM) method. For this purpose we use Simple additive weighting method (SAW). In SAW, the problem can be expressed in decision matrix D as in (3.2), where row i corresponds to alternative link N_i and column j corresponds to type of link layer metric.

$$D = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} delay_{wifi} & loss_{wifi} \\ delay_{wimax} & loss_{wimax} \end{bmatrix} \quad (3.2)$$

$$W = \begin{bmatrix} w_{delay} & w_{loss} \end{bmatrix} \quad (3.3)$$

$$x'_{ij} = x_j^{min} / x_{ij} \quad i = 1, 2 \quad j = 1, 2 \quad (3.4)$$

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

We defined W as weight matrix (3.3), where w is the weight of each criterion. As delay and loss metric are categorized as cost criteria, we normalize each element in the decision matrix D using (3.4) where x_{ij} is the value of alternative N_i with respect to criteria x_j . Thereafter we apply the weight matrix to the normalized decision matrix. Our VDA will choose the interface which has a greater merit value to send out the packet.

Figure 3.11: Policy impact on application's throughput

We evaluate the performance of adaptive load balancing through simulation using ns-2 network simulator [35]. The performance metrics used are application's throughput, end to end delay, and inter-packet delay variation [36]. The simulation result is shown in Fig. 3.11 to Fig. 3.13. We use exponential traffic generator to simulate VoIP traffic. We set both w_{delay} and w_{loss} to 0.5 to indicate equal importance. We simulate an environment where increasing number of mobile nodes are attached to both WLAN AP and WMAN BS while flooding the networks with packets transmission. In Fig. 3.11, when the number of mobile nodes increases beyond 50 nodes,

CHAPTER 3. MIH BASED MULTI-HOMED SIP MOBILITY MANAGEMENT (MHMM)

Figure 3.12: Policy impact on application's end to end delay

the voice throughput in single interface scheme drops drastically. The result shows that employing adaptive load balancing scheme improves data throughput and reduces average end to end delay as well as inter-packet delay variation, especially under congested network. Utilizing only single interface for the whole duration of traffic flow is disadvantageous as it results in degradation of user's experience since the access network QoS is not considered. Adaptive load balancing policy overcomes this problem by spreading the traffic flow while monitoring the parameter change at the link layer. These new averaged parameter values are always taken into account when VDA decides the network interface to send the packet out at every interval. The scheme enables terminal to predict the access network performance dynamically and load balance the traffic based on this information. As a result, delay sensitive application is better supported.

Figure 3.13: Policy impact on application's inter-packet delay variation

3.5 Summary

In this chapter, we have presented MIH-based Multi-homed SIP mobility management scheme (MHMM) and illustrated its performance via a 2-interface (Wi-Fi, WiMAX) terminal. We highlight one key component, i.e. VDA, for its importance in maintaining the application's connection state by ensuring IP address transparency, supporting multiple addressed of the different interfaces as well as providing intelligent operation in accordance to the policies. Our scheme, with the aid of MIH Link Event trigger, enables seamless communication during vertical handover. For optimizing interface usage, our scheme exploits the MIH framework to obtain dynamic link layer information from underlying interfaces and use the information as metrics in the policy. We introduce adaptive load balancing policy which is tailored especially for delay sensitive application. The scheme incorporates features of both packet-based and flow-based load balancing and tries to minimize out-of-order packet problem. The result shows that the scheme works effectively in rapid changes of networks' performance.

Chapter 4

Adaptive Multiple-Interface Management Framework for Multihomed Terminal (AMIMF)

This chapter extends MIH Based Multi-homed SIP Mobility Management (MHMM) described in Chapter 3 to provide an adaptive and environment aware framework to enhance user's perceived quality.

As mentioned earlier, multihoming capability, supported by a device having more than one network interface, provides several benefits such as load balancing between network interfaces, redundancy when one network interface fails, and ubiquitous access through multiple service providers. The new framework, called Adaptive Multiple-Interface Management Framework for Multihomed Terminal (AMIMF) is aware of the network environment and adopts different scheduling schemes and policies in accordance with the changing network conditions for efficient delivery of data.

An overview of AMIMF is shown in Fig. 4.1 using a 3-interface terminal device as an illustration. The terminal device is configured with three network interfaces of diverse technologies, i.e. IEEE 802.11 (Wi-Fi), IEEE 802.16 (WiMAX), and cellular UMTS. Having this configuration, we benefit from better coverage provided by

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

UMTS and WiMAX and aggregate capacity of each access technology. When handling traffic, the terminal may switch alternatively among its three interfaces or use these interfaces simultaneously depending on the policy residing in the terminal. In the context of multi-access for terminal, policy represents preferences and guidance on how packet will be forwarded to the destination. We limit the scope of this chapter to demultiplexing the traffic flow based on user's perspective with multihomed terminal as the source.

Three main functional blocks are defined in the architecture, namely VDA entity, MIH entity, and Profile entity¹.

In AMIMF, we introduce a new agent called profile agent. This agent, together with profile storage, forms the Profile entity. Profile storage stores L2 parameters measured from all network interfaces. As we add UMTS interface to the device, we include available UMTS bandwidth parameter to be saved in the profile storage. Computing available bandwidth in UMTS is straightforward as the communication session is connection-oriented. Awareness of access network performance is important when sending data packet such that the terminal may maximize its network connections. This is achieved through an operation called access network performance prediction performed by the profile agent. Profile agent retrieves link layer information from storage, especially of Wi-Fi and WiMAX interface, and determines the performance quality of access networks. The result is stored back in the profile storage and used by VDA to help execute the policy.

4.1 Network Interface Selection

Interface selection is handled by VDA to determine the local interfaces used to send out packets. In the case of multihomed terminal, there could be more than one end-to-end path from terminal to destination. Thus further processing is required

¹Refer to Section 3.1 of Chapter 3 for more details

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

Figure 4.1: Architecture of multihomed terminal equipped with three interfaces: Wi-Fi, WiMAX, and UMTS

to determine which interface yields the greatest benefit for the user. Choosing the network interface in an arbitrary manner could be detrimental. Thus, QoS factors depicting the access network condition should be taken into account in the interface decision process. These factors constitute link layer measurements. Since this measurement is dependent on the type of access technology, it is necessary for an efficient framework to make them deliverable to upper layer protocols. This framework facilitates cross-layer information exchange between lower and upper layers.

Recently, a standard called MIH has emerged and is being developed by the IEEE 802.21. This standard suits the proposed AMIMF due to the fact that it is developed to be aware of available diverse radio access technologies. MIH provides link layer intelligence to upper layer protocols. The main objective of MIH is to sup-

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

port handover between different radio technologies, referred to as vertical handover. Cross-layer information exchange is facilitated by an entity called MIH function which resides between link layer and network layer. However, MIH has some limitations. According to the standard, event trigger and measurement gathered at link layer are used solely to optimize handover. The event will be triggered if the measurement crosses certain preset threshold. In our scheme, we extend the operation of MIH. MIH will not only support handover as its primary objective, but also help execute the policy in the terminal. For this purpose, MIH_Link_Parameters_Report primitive is used to convey link layer information regularly to the upper layers.

The wireless access network portion may contribute to significant performance bottleneck in the overall network performance due to its dynamic nature. We rely on MIH framework to capture the network dynamic. The existence of many terminals within a single collision domain of wireless access network, especially WLAN IEEE 802.11 may cause congestion which in turn contributes to larger end-to-end delay, dropped packets, etc. Thus in our policy, it is necessary to take into account the condition of access network when selecting the interface. This includes consideration of congestion in the access network.

Bejerano et al. [37] use channel busy time to represent the load metric in wireless LAN. Channel busy time usage is twofold, one is to represent the load on the network and another is to indicate the level of congestion the access point experiences. Jardosh et al. [38] study congested wireless LAN network empirically and draw the correlation between network congestion and its effect on link layer performance. In our scheme, we also use channel busy time which is represented by the portion of time that the station senses the medium as busy. To represent congestion in WiMAX network, we utilize a link layer metric called Bandwidth Request (BW-REQ) packet loss ratio. In addition to the two previously mentioned metrics, we also use L2 delay to better predict the access network performance. Further explanation of these metrics will be given later. In this chapter, we show how the congestion, as a function of L2

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

delay and channel busy time (in WLAN) as well as L2 delay and BW-REQ loss ratio (in WMAN), impacts the measured throughput of a network. Based on this result, we subsequently determine a set of thresholds indicating the level of access network performance. Thus, our work relates the access network performance level to the estimated available throughput that the network can provide to support the application.

Many works have been done regarding available bandwidth estimation which can be categorized into two main categories: intrusive and non-intrusive scheme. Intrusive scheme, such as in [39] [40], uses end-to-end probe packet to estimate path bandwidth. Using intrusive measurement poses a drawback that it may cause large overhead and waste precious wireless network resource, especially under the condition where many multihomed terminals in the network need to estimate bandwidth through all links. On the other hand, non-intrusive scheme utilizes the trace history of previous data transmission.

In this chapter, we propose an idea to equip the terminal with prior gleaned knowledge from data packets to help terminal estimate available throughput in the access network. Having this data, terminal, especially profile agent, can predict the access network performance level. For this purpose we conduct offline simulations to infer the correlation between measured load, channel busy time (for WLAN) or BW-REQ loss ratio (For WMAN), and L2 delay. Based on the results, we classify the performance of access network into three categories and determine the thresholds associated to them which serves as prior knowledge to the terminal.

4.2 Cross Layer Information

We extract some information from the MAC layer of Wi-Fi and WiMAX interface and send the information to upper layer (i.e. VDA) using the MIH framework. The information are as follows:

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR
MULTIHOMED TERMINAL (AMIMF)

4.2.1 L2 Delay

Wi-Fi, in comparison with WiMAX, has different mechanism in the channel access for the data transmission. In Wi-Fi, the channel access for the data transmission is based on contention and the station relies on the acknowledgement scheme to indicate successful transmission of the data frames. In contrast, the channel access in WiMAX is based on the granted access in which the base station (BS) allocates uplink (UL) burst for the station, thus avoids the possibility of collision during the data transmission period and eliminates the need for an acknowledgement scheme. As the MAC of both Wi-Fi and WiMAX operates under different mechanisms, we define the L2 delay term separately. Denote T_{queue} as the time when packet is ready in the queue for transmission; T_{ACK} as the time the terminal receives packet acknowledgement (ACK); $T_{transmit}$ as the time the terminal actually transmits the packet; and tx_time as the time needed to transmit the whole packet. L2 delay of a packet i in Wi-Fi network is calculated using Eq. 4.1, whereas L2 delay of packet i in WiMAX network is calculated using Eq. 4.2.

$$D_{(i)}^{Wi-Fi} = T_{ACK(i)} - T_{queue(i)} \quad (4.1)$$

$$D_{(i)}^{WiMAX} = (T_{transmit(i)} - T_{queue(i)}) + tx_time(i) \quad (4.2)$$

We perform similar simulations for both Wi-Fi and WiMAX. Firstly there is a terminal sending traffic to a receiver through the network. We then measure the L2 delay at the sender terminal. To know the impact of network load and contention on the value being measured as a function of node number, we continually increase the background traffic by adding terminals that transmit packets through the network. The simulation is performed using terminal number starting from 1 to 15. Fig. 4.2 shows the relation between L2 delay and normalized traffic load in Wi-Fi network. The normalized load is obtained by dividing the measured load with the

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

transmission rate, Tx_rate . In our simulation, Tx_rate for Wi-Fi is 11 Mbps. From the observation, as the load and the number of terminals increase in the network, L2 delay also increases. For WiMAX, the result is shown in Fig. 4.3. The same trend is observed, but WiMAX is less sensitive than Wi-Fi in terms of increase in L2 delay corresponding to an increase in network load. Note that for WiMAX, Tx_rate used for normalizing the load is 7,962,352 bps.

Figure 4.2: Normalized load vs L2 delay in Wi-Fi

4.2.2 Channel Busy Time

A network link overwhelmed with too much traffic causes congestion. Consequently, the terminal experiences network throughput degradation. To help indicate congestion in Wi-Fi network, we use channel busy time metric obtained from the Wi-Fi interface. Channel busy time (T_{total}^{busy}), comprises two main parts: busy transmitting T_{tx}^{busy} and busy receiving time T_{rx}^{busy} . This metric is the summation of all transmitted

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

Figure 4.3: Normalized load vs L2 delay in WiMAX

and received frames on the terminal's Wi-Fi MAC in one period T_{int} . The measurement granularity is set at T_{int} equals 1 second.

$$T_{tot(t)}^{busy} = \frac{T_{tx(t)}^{busy} + T_{rx(t)}^{busy}}{T_{int}} \quad (4.3)$$

The channel busy time ratio, $T_{tot(t)}^{busy}$ is computed according to Eq. 4.3, with suffix (t) indicating the measurement in the "t-th" period. The computed value represents the fraction of time interval that the channel status is busy. For the measurement, we consider all kinds of packets including data, management (Beacon), and control packets (RTS, CTS, ACK). The basic access method of IEEE 802.11 is Distributed Coordination Function (DCF) protocol. DCF employs carrier sense multiple access with collision avoidance (CSMA/CA) for channel access. When the terminal wants to transmit a frame, the terminal has to sense that the medium is idle for a duration of DIFS (Distributed coordination function Interframe Space). This period is followed with a Backoff interval (BO). If the medium is idle after BO expires, terminal proceeds

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

with transmitting the frame to the destination. The destination, upon receiving the frame, sends ACK back to the terminal. RTS-CTS mechanism is enabled in the operation to further reduce the collision probability and avoid hidden terminal problem. Retransmission time as a result of error delivery is also taken into account in the busy time calculation.

Fig. 4.4 shows the relation between channel busy time and normalized load in the WLAN network. We continually increase the load and the number of terminals by adding terminals that transmit packets through the network, starting from 1 to 15 terminals. As the load and the number of terminals increase in the network, channel occupancy also increases. This is mainly due to collision and retransmission. From the figure, we can observe a period when the measured load decreases after reaching a certain point (this point is the saturation capacity of the WLAN network). In our opinion, the wireless channel is highly congested during this period.

Figure 4.4: Normalized load vs channel busy time in Wi-Fi

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR
MULTIHOMED TERMINAL (AMIMF)

4.2.3 Bandwidth Request (BW-REQ) Loss Ratio

In point-to-multipoint (PMP) mode of IEEE 802.16, a base station (BS) serves a number of subscriber station (SS) [41] [32]. In our work, we consider the network operates in a time division duplex (TDD) where a frame is divided into uplink (UL) and downlink (DL) subframes. As part of QoS support, a bandwidth request-grant mechanism is applied according to the standard to reduce the data collision. In Best Effort (BE) traffic class, for uplink scheduling, SS first has to request for bandwidth individually from its serving BS every time it has packet waiting in the connection queue. This is done by sending a BW-REQ packet via the shared BW-REQ slots in UL subframe. Since there is a limited number of slots, SS has to compete with other SS which denotes the start of the contention resolution process (CRP). SS subsequently starts a backoff process with initial window equal to W_0 . The SS shall randomly select a number within interval $[0, W_0 + 1]$ as the backoff interval. When the backoff expires, SS sends BW-REQ packet. If transmission is successful, upon receiving the request, BS may either approve and allocate slot for SS's data transmission or reject the request if there is not enough resource. CRP is completed if SS receives the grant within the period T_{REQ} . On the contrary, if SS does not receive grant within the period T_{REQ} , SS concludes that BW-REQ transmission has failed, thus starting a new backoff and increasing the backoff window by a factor of two. The new backoff window will be $2^i W_0$, where i is the i_{th} retry of BW-REQ transmission, and backoff is chosen randomly within interval $[0, W_i + 1]$. If the maximum number of retries, m (which is set to 3), has been reached, and BW-REQ packet transmission still fails, SS discards the packet and CRP ends. As the bandwidth request-grant mechanism in BE service is contention based, this scheme works fine if the channel is not congested (when there is enough BW-REQ slots for all SSs' connections). However there is significant performance degradation in the case of heavy load in the network as well as excessive number of SS requesting for bandwidth. When the load is high, the queue in each SS is never empty. Thus if there is more than one SS trying to transmit

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

their request via a slot, collision is likely to occur and the bandwidth request packets collide and the SS does not receive the allocation. It may cause the SS to suffer a longer delay and reduced throughput.

Figure 4.5: Normalized load vs BW-REQ loss ratio in WiMAX

$$BW_REQ_{loss(t)} = \frac{N_{corrupt(t)}^{BW_REQ}}{N_{sent(t)}^{BW_REQ}} \quad (4.4)$$

Bandwidth request packet loss ratio BW_REQ_{loss} at measurement interval t , is computed using Eq. 4.4. $N_{sent(t)}^{BW_REQ}$ includes all transmission and re-transmission attempts and $N_{corrupt(t)}^{BW_REQ}$ is incremented by 1 every time T_{REQ} timer expires without SS receiving the bandwidth allocation. Fig. 4.5 shows the relation between BW-REQ loss and normalized network load. In this simulation, we continually increase the load in the network by adding terminals (from 1 - 15 terminals) while sending background traffic through the network. As seen in Fig. 4.5, as the number of SS and the load increase, collision occurs more frequently and contributes to greater

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR
MULTIHOMED TERMINAL (AMIMF)

bandwidth request packet loss ratio.

Table 4.1: Profile on access network performance

Access network	State	Normalized Load (L_{norm})	L2 delay (D msec)	Busy time ratio (T^{busy})
IEEE 802.11 (3.5 Mbps)	State 1	$0 \leq L_{norm} \leq 0.143$	$D \leq 5.378$	$T^{busy} \leq 0.391$
	State 2	$0.143 < L_{norm} \leq 0.285$	$5.378 < D \leq 9.867$	$0.391 < T^{busy} \leq 0.783$
	State 3	$L_{norm} > 0.285$	$D > 9.867$	$T^{busy} > 0.783$

Access network	State	Normalized Load (L_{norm})	L2 delay (D msec)	BW Req loss ratio ($Loss$)
IEEE 802.16 (5 Mbps)	State 1	$0 \leq L_{norm} \leq 0.194$	$D \leq 140.850$	$Loss \leq 0.356$
	State 2	$0.194 < L_{norm} \leq 0.555$	$140.850 < D \leq 229.794$	$0.356 < Loss \leq 0.445$
	State 3	$L_{norm} > 0.555$	$D > 229.794$	$Loss > 0.445$

4.3 Access Network Performance State

Having all the abovementioned results, we notice that the observed throughput has a close inter-dependancy with the level of congestion in the access network. Therefore we suggest that the access network be classified into a set of performance categories according to its congestion state. This is achieved by observing the trend in network load with respect to increasing {channel busy time, bandwidth packet request loss ratio} and L2 delay in {Wi-Fi, WiMAX} networks. As the sensitivity of Wi-Fi and WiMAX access network towards congestion differs, with Wi-Fi being more sensitive than WiMAX, we define different range of performance category for both access networks. With reference to the measurements done, we classify the access network

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR
MULTIHOMED TERMINAL (AMIMF)

Table 4.2: Estimated load and available throughput in our framework

Access network	State	Est. load range (L kbps)	Est. av. throughput (B_{av} kbps)
IEEE 802.11 (3.5 Mbps)	State 1	$0 \leq L \leq 1,568$	$1,932 \leq B_{av} \leq 3,500$
	State 2	$1,568 < L \leq 3,136$	$364 \leq B_{av} < 1,932$
	State 3	$L > 3,136$	$B_{av} < 364$
IEEE 802.16 (5 Mbps)	State 1	$0 \leq L \leq 1,548$	$3,452 \leq B_{av} \leq 5,000$
	State 2	$1,548 < L \leq 4,419$	$581 \leq B_{av} < 3,452$
	State 3	$L > 4,419$	$B_{av} < 581$

into 3 states as shown in Table 4.1 with the following interpretation:

- State 1: Uncongested condition
- State 2: Moderately congested condition
- State 3: Highly congested condition

Note that 3.5 Mbps and 5 Mbps denote the capacity of Wi-Fi and WiMAX respectively. These values are obtained from the saturation throughput we obtained through simulation. Assuming that at certain time, the terminal knows the total capacity of certain access network and given that it also has knowledge of the network load for a known access network state, the terminal may compute an estimation of available throughput of the network at that time instant. This information could be useful for terminal when it decides the interface to send out the packets, especially if the user's satisfaction level is focused on delivery throughput of the network. Specific to Wi-Fi and WiMAX technologies we use in our simulation, we devise estimates of load and available throughput as shown in Table 4.2. For example, based on the

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

table, we say that a State 2 Wi-Fi access network can provide a range of throughput from 364 kbps to 1,932 bps to support terminal's data transmission.

To determine the performance state of access network, Profile agent retrieves parameters D^{Wi-Fi} , D^{WiMAX} , T_{tot}^{busy} and BW_REQ_{loss} from storage and then compare the values against the specification in Table 4.1. This process recurs at time interval of 1 second.

4.4 Packet Distribution Policy

Packet distribution scheme determines the level of satisfaction the users achieve when using the network service. In [42] we introduce a scheme called adaptive load balancing. In this scheme, VDA takes into account the network condition and adaptively switches between interfaces to send out the packet. Where users have capability to access more than one network at a time, applying policy to balance the traffic among available network interfaces is beneficial. This balancing strategy is applied in conformance to resource aggregation objective of having multiple network interfaces in a single terminal. It is beneficial in a scenario where traffic flow QoS requirement cannot be fully satisfied by network resource provided through a single interface. By splitting the traffic flow into multiple sub-flows and transmitting them simultaneously over different sets of network interfaces, user perceived quality can be enhanced as more throughput is available. VDA takes responsibility to implement this strategy when it receives packets coming from the upper layer. VDA processes the packets in accordance with the policy defined in the system for a certain flow.

We compare several load balancing schemes and present the performance result.

4.4.1 Round Robin (RR)

In RR scheme, VDA transmits packet alternatively to each network interface in order and in equal allocation. For this purpose, VDA only needs to keep track of up and

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

ready interfaces which provide end-to-end path to the destination. The terminal does not necessarily monitor the dynamic nature of the link quality.

4.4.2 Adaptive Distribution

In this scheme, packet by default will be sent through the network in the following order: State 1, State 2, State 3 in Tables 4.1 and 4.2. UMTS interface is excluded. If both Wi-Fi and WiMAX access network fall into the same classification such as both in State 1 or 2, the Wi-Fi interface will take precedence over the WiMAX interface. If both access networks are classified as State 3 network, packet will be sent through UMTS interface. Under low and moderately saturation conditions, Wi-Fi network is prioritized over WiMAX as it is cheaper in terms of usage cost.

4.4.3 Adaptive Weighted Round Robin (Adaptive WRR)

Table 4.3: Throughput value used for adaptive WRR

Access network	State	Throughput (R_{av} kbps)
IEEE 802.11	State 1	3,500
	State 2	1,932
	State 3	364
IEEE 802.16	State 1	5,000
	State 2	3,452
	State 3	581

In this scheme, we implement WRR scheduling with adaptive weight. The weight for network interfaces is derived from the ratio of estimated available throughput provided by each access network. As we classify the Wi-Fi and WiMAX network performance in terms of provided throughput range, we determine a set of discrete throughput value to be used for the weight calculation. This is not an issue with UMTS interface as we can know the exact available throughput at any time given

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR
MULTIHOMED TERMINAL (AMIMF)

that connection is guaranteed with certain allocation of bandwidth. We use the throughput value from Table 4.3. If L is a set of access networks on which a terminal is connected; R_{av} is a set of throughput provided by the access network in L ; the WRR weight for each interface w_i is calculated using Eq. 4.5. Since the offered throughput changes when access network falls into different State, the weight assigned to each interface also changes over time. The packet distribution follows in accordance with the weights.

$$w_i = \frac{R_{av_i}}{\text{Min}(R_{av_i}; \forall i \in L)} \quad (4.5)$$

4.4.4 Hybrid Scheme

This scheme is similar to Adaptive WRR. The only difference is when Wi-Fi access network is classified as State 3, Wi-Fi interface will be neglected for the interface selection and VDA will choose the access network having the smallest State number. Other than that condition, adaptive WRR scheme will be used. When Wi-Fi network falls into State 3, the access network is heavily congested. As we have seen, Wi-Fi network is more sensitive towards congestion than WiMAX network and its performance degrades drastically. Thus considering Wi-Fi interface as non-eligible interface for transmission at this state may lead to better performance.

4.5 Simulation and Results

Fig. 4.7(a) shows the result of end-to-end packet delay measurement. We conduct simulations using ns-2 [35] in order to assess the performance of load balancing policies we have outlined previously. The performance metrics used are achieved throughput, number of packets received, end-to-end packet delay, and inter-packet delay variation [36]. Inter-packet delay variation of a packet is defined as end-to-end delay difference of a packet with regards to the previous packet in the stream. In our measurement,

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

for example, if within the measurement interval $[0 - t]$ we receive packet i with $i = 1, 2, 3, \dots, n$, we take the average value of inter-packet delay variation, $IPDV$, which is given by $\overline{IPDV} = \frac{\sum_{i=1}^n IPDV_{(i)}}{n-1}$, where $IPDV_{(i)} = |Delay_{(i)} - Delay_{(i-1)}|$.

The measurement for all performance metrics is done with regards to all packets originating from multihomed terminal. Our simulation scenario is as follows: The multihomed terminal is located in the coverage of UMTS, Wi-Fi, and WiMAX access network. The UMTS connection is allocated with an uplink dedicated bandwidth of 384 kbps. At $t = 15$ s, terminal starts sending Constant Bit Rate (CBR) traffic to the destination node which is around 2 Mbps in total. To slowly saturate the network, we add terminals sending background traffic of 384 kbps one at a time. For each interval $\Delta t_1 = 10$ s, starting from $t = 10$ s, we introduce a single interface terminal connecting to Wi-Fi network. Therefore at $t = 20, 30, 40, \dots, 170$ s, a Wi-Fi terminal is added, giving a total number of 16 terminals with Wi-Fi interface in the simulation environment. In the mean time, for each interval $\Delta t_2 = 10$ s, starting from $t = 15$ s, we also introduce a single interface terminal connecting to WiMAX network. Therefore at $t = 25, 35, 45, \dots, 175$ s, a WiMAX terminal is added, giving a total number of 16 terminals with WiMAX interface in the simulation environment. Thus, after $t = 175$ s, there are 33 terminals in the network, including the multihomed one. The simulation is run until $t = 300$ s and the result is presented in Fig. 4.6 and Fig. 4.7.

As shown in Fig. 4.6(a), Hybrid scheme outperforms other schemes in that it achieves more throughput, whereas RR scheme shows the poorest performance. This is because in RR scheme, VDA tries to assign traffic equally among all interfaces without considering the access network conditions and available capacity. From $t = 15$ s to 110 s (up to $t = 110$ s, 9 Wi-Fi and 9 WiMAX terminals are in the network), Hybrid scheme performs similarly to Adaptive WRR scheme, and thereafter, Hybrid scheme outperforms Adaptive WRR scheme. This is because during the time duration up to 110 s, access network starts moving from uncongested to moderately congested

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

state. Under these states, access networks, in this case Wi-Fi and WiMAX networks, still manage to support the applications. However for $t > 110$ s, both networks start to approach the highly congested state (classified as State 3), with Wi-Fi changes state faster as it is more sensitive. Since the Wi-Fi network under high congestion state shows more significant performance degradation compared to WiMAX, using Wi-Fi interface is disadvantageous. Under Hybrid scheme, being aware of this condition, VDA marks the Wi-Fi interface as non-eligible and proceeds to select interface with the best state, i.e. WiMAX interface. In terms of received packet number, as shown in Fig. 4.6(b), Hybrid scheme performs the best which means this scheme has less dropped packets compared to other schemes. Using RR scheme incurs many dropped packets. The result from received packet number are consistent with that of achieved throughput.

Fig. 4.7(a) shows the Hybrid scheme and Adaptive Distribution scheme are relatively better than the other two schemes in terms of end-to-end packet delay. Given that end-to-end packet delay, T_{e2e} , is the summation of packet delay in access network, T_{an} and packet delay in core network, T_{core} . Assuming that the core network performance is stable due to static and robust infrastructure, T_{core} does not change significantly. Hence T_{e2e} mainly depends on T_{an} . Thus, end-to-end packet delay value is mainly contributed by the delay from terminal to AP or BS. The biggest portion contributing to higher end-to-end delay resulting from RR and Adaptive WRR scheme comes from packets which are transmitted through the Wi-Fi network. While these two schemes constantly use Wi-Fi interface for transmission under all congestion states of network (represented by the network state), Hybrid scheme and Adaptive Distribution scheme prefer not to use Wi-Fi interface when the Wi-Fi network is highly congested. For the same reason, the latter schemes hence produce better result in inter-packet delay variation performance as seen in Fig. 4.7(b). The fluctuation of inter-packet delay variation resulting from RR and Adaptive WRR scheme is contributed by the inter-packet delay variation from the packets which are received

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

through the Wi-Fi network.

4.6 Summary

In this chapter, we propose an adaptive multi-interface framework, AMIMF, to enhance the user perceived quality. The framework obtains MIH metrics, e.g. L2 delay, channel busy time ratio, bandwidth request loss ratio, and uses the observed values to infer access network performance according to defined classification. We have shown that the classification also relates to the congestion level that the access network experiences. We also compare several load balancing schemes, such as RR, Adaptive WRR, Adaptive Distribution, and Hybrid scheme. Simulation result shows that Hybrid scheme performance outperforms other schemes in terms of throughput, end-to-end delay, inter-packet delay variation, and packet drop, while RR scheme shows the worst result. Based on the results shown, we also suggest that it is not feasible for the terminal to use Wi-Fi interface when Wi-Fi network is highly congested and needs to deem the interface non-eligible for packet transmission. It thus shows that our proposed AMIMF exploits the terminal awareness of the access network condition to greatly enhance user's perceived quality, in particular for application with long connection duration, such as media streaming.

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

(a)

(b)

Figure 4.6: Policy impact on (a) throughput ; (b) number of packets received

CHAPTER 4. ADAPTIVE MULTIPLE-INTERFACE MANAGEMENT FRAMEWORK FOR MULTIHOMED TERMINAL (AMIMF)

(a)

(b)

Figure 4.7: Policy impact on (a) end-to-end delay ; (b) inter-packet delay variation

Chapter 5

QoS Support in Multihomed Terminal

In this chapter, our Adaptive Multiple-Interface Management Framework for Multihomed Terminal (AMIMF) is further integrated with QoS support of underlying access technologies to better manage application's QoS requirement. Different types of applications have different needs in terms of latency, bandwidth, packet error rate, etc. Consequently, QoS support to cater for all types of application is required. While various access network technologies exhibit different characteristics in terms of coverage, delay, and bandwidth, the ability to maintain user's perceived QoS across these technologies becomes important. This is further complicated by the growth of mobile users who frequently roam from one network to another as well as the demand of multimedia application. This chapter presents the integration of our Adaptive Multiple-Interface Management Framework (AMIMF) with L2 QoS mechanisms, of both Wi-Fi and WiMAX, to better support various applications over the wireless links. We limit our discussion to QoS handling over the wireless access networks.

Wu et al. in [43] compare the performance of MADM-based handover decision scheme using MIH with the received signal strength (RSS) based and cost function based vertical handover decision scheme. However the work only uses RSS as the trig-

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

ger for handover, while our framework exploits more L2 metrics as handover triggers to optimize user's access.

5.1 QoS in WLAN

Figure 5.1: Access Categories in EDCA mode

The original IEEE 802.11 WLAN standard [44] has been initially designed without QoS support, aiming at providing best effort at low cost. Hence, the protocol provides same treatment for all kinds of applications. IEEE 802.11e [45] proposes QoS enhancement in the MAC layer. The QoS is provided by differentiating the application based on its type. The IEEE 802.11e standard introduces Hybrid Coordination Function (HCF) which operates in two modes, ie. Enhanced Distributed Coordinated Access (EDCA), which is contention based and HCF Controlled Channel Access (HCCA) which is polling based. HCCA is not widely implemented (the reason can be found in [46]). EDCA improves DCF mode in the original standard and de-

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

defines four Access Category (AC) existing in the MAC layer, namely AC_VO (for voice traffic), AC_VI (for video traffic), AC_BE (for best effort traffic), and AC_BK (for background traffic) ordered from the highest to lowest priority. Figure 5.1 depicts four ACs for EDCA. Each AC has its own transmission queue and associated parameters. To achieve prioritization among ACs, the following parameters are configured:

- Minimum contention window (CW_{Min}): minimum value of random backoff
- Maximum contention window (CW_{Max}): maximum value of random backoff. AC is first initialized with CW_{Max} . Random backoff window size is doubled upon each unsuccessful attempt of channel access with upper bound of CW_{Max} .
- Arbitrary Inter-Frame Spacing (AIFS): interval of time during which a station has to wait after detecting that the channel is idle and starts its backoff timer.
- Transmission Opportunity (TxOP) limit: with the TxOP burst option being enabled, station is allowed to send a number of consecutive packets without intermediate contention upon gaining access to the channel. TxOP limit specifies the maximum duration of the data frame burst. However, for simplicity, we disable this option during simulation.

Higher priority AC is usually assigned lower CW_{Min} , CW_{Max} , and AIFS values than the lower priority AC. This ensures time-critical traffic has higher probability to occupy the channel. To handle virtual collision among internal ACs, virtual collision resolution is applied where the AC with the highest priority takes precedence to access the channel, while the remaining ACs double their backoff window size in the event of internal collision.

$$AIFS[AC] = SIFS + (AIFSN[AC] * SlotTime) \geq DIFS \quad (5.1)$$

The value of AIFS for each AC ($AIFS[AC]$) is given by Eq. 5.1, where SIFS is Short Inter-frame Spacing with duration 10 μs , AIFS number ($AIFSN$) is an integer greater

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

than zero, and $SlotTime$ is $20 \mu s$. The equation denotes that the AIFS assigned to an AC is at least equal to the DIFS used in the legacy MAC. The setting of DIFS follows Eq. 5.2 which yields $50 \mu s$.

$$DIFS = SIFS + 2 * SlotTime \quad (5.2)$$

5.2 QoS in WMAN

Different from IEEE 802.11, IEEE 802.16 WMAN MAC [47] is connection-oriented. All traffic, either in uplink or downlink direction, are carried using dedicated connections. To facilitate traffic flow with different set of QoS requirement, IEEE 802.16 standard defines the following four service classes (SC) which are detailed in Table 5.1: Unsolicited Grant Service (UGS), Real-Time Polling Service (rtPS), non-Real-Time Polling Service (nrtPS), and Best Effort (BE). Five QoS parameters for traffic flow resource reservation are used, i.e. maximum delay, jitter, minimum and maximum data rate, traffic handling priority. To classify the frame into already established connection, connection identifier (CID) is included in the frame header.

Table 5.1: IEEE 802.16 service class

Service class	Description
UGS	supports CBR traffic, such as VoIP without silence suppression
rtPS	supports real-time service with variable-size packet on a periodic basis, such as VoIP with silence suppression and MPEG
nrtPS	supports non-real-time service which requires variable size of burst granted on a regular basis, such as FTP
BE	supports application that does not require explicit QoS support, such as HTTP Web

5.3 Inter-working Mechanism in QoS-enabled Terminal

Figure 5.2: Illustration of WLAN and WMAN network integration

Interworking mechanism for user to roam across the networks depends highly on the integration architecture. In our proposal, the architecture for integration of WLAN (Wi-Fi) and WMAN (WiMAX) access networks is *loose-coupling*, as illustrated in Figure 5.2. In this type of architecture, WLAN and WMAN access networks are completely separated and both networks are connected via the core network. This allows the scenario where access network is deployed independently

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

and may belong to different network operators. Terminal-controlled handover usually applies under this scheme. As the user has connectivity support via diverse access technologies, possibility of roaming or utilizing all network interfaces simultaneously becomes an important consideration. Increasing popularity of time-critical applications drives the need for seamless handover to guarantee session continuity. However, as discussed before, the difference in QoS mechanism, implemented in IEEE 802.11e and IEEE 802.16 technologies, exposes a challenge. Application may experience QoS degradation after handover to another access network of different technology. It necessitates a requirement for a QoS mapping strategy between IEEE 802.11e and IEEE 802.16 technologies to ensure QoS consistency across network domains. Some works in [26] [27] discuss the QoS mapping issue during handover to support seamless service continuity. A scenario when user roams from WLAN to WMAN network is also presented in [26].

As the interfaces in a terminal are of specific technologies, an intelligent framework is needed to manage and coordinate the multi-interface of the terminal efficiently. By exploiting L2 metrics which has been discussed before, the framework is thus rendered environmentally-aware of the prevailing access networks conditions. The link layer observation is used further to aid management-related tasks, such as network interface selection and application handovers with the ultimate goal to enhance user perceptive quality. RSS, which is used in conventional handover decision, cannot stand as the sole factor to indicate wireless network performance. However, the metric is useful as a trigger for terminal to initiate handover, especially due to terminal mobility.

Note that even when mobility is introduced as exemplified by a scenario in Figure 5.2: the user goes out of WLAN 1 cell, crosses into the WMAN coverage, and then enters into WLAN 2 cell, our framework can still maintain the QoS and ensure seamless handover. This is accomplished via the multi-interface capability of the terminal. The terminal employs *make-before-break* concept whereby when one interface (Wi-Fi) is used for packet transmission, the other interface (WiMAX) is used

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

for signaling purpose in handover preparation. The traffic is then routed through the WiMAX while the terminal exits from one WLAN cell. As the terminal enters the next WLAN cell, the Wi-Fi interface will make a new connection with the new WLAN cell and the traffic is then routed from the WiMAX to the new WLAN. Hence a seamless handover is achieved with no break in the traffic transmission.

In addition, our framework becomes more intelligent by exploiting channel busy ratio and bandwidth request loss ratio metrics to cope with the handover decision. We explain our scheme through the following scenario. Terminal concurrently initiates three different applications to CN, i.e. voice, video streaming, and data transfer. Upon receiving packets from the upper layer, VDA extracts information from the packet header and checks its connection table. If there is already an established connection for the corresponding traffic flow, VDA retrieves the network interface in association with the flow and use it as the sending interface. If not, network interface selection is initiated. Note that the flow-interface association may change over time as handover occurs. VDA utilizes *flow-based* scheduling to handle all types of applications in the current scheme. During association establishment of the applications, VDA prioritizes Wi-Fi over WiMAX interface as the licence-exempt Wi-Fi bandwidth incurs a much lower cost if not free. This preference benefits user in acquiring good QoS at the lowest possible cost. However it requires further checking of the WLAN network to ensure support for the applications. If the terminal senses that its serving WLAN network is not in saturated condition (as implied by its channel busy ratio parameter), Wi-Fi interface is selected for the application. The association between the application and the Wi-Fi interface is then inserted into a connection table maintained by the VDA. VDA employs IP swapping method [12] [13] to avoid session disruption due to IP address change during vertical handover. The IP address of each packet is swapped at the network layer to reflect the current new IP addresses and subsequently restored so that the change in IP address which takes place during handovers and mobility is transparent to the applications. Voice, video streaming,

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

and data transfer applications are classified into AC_VO, AC_VI, and AC_BE respectively when they are serviced in the WLAN network.

5.3.1 WLAN $\rightarrow$ WMAN handover

Two envisioned handover policies are currently implemented in the terminal based on user preferences, namely Policy I and Policy II. Users can indicate their preferences any time during the session. However, the number of policies could be extended to cater to the diverse user's preferences. Under Policy I, VDA handovers all applications to the WMAN network for simplicity purpose whereas under Policy II, VDA prioritizes its critical-time applications (e.g. voice and video streaming application) to be handed over first. Data application remains at WLAN network while terminal continues to monitor its channel busy time ratio. If the channel busy time ratio still exceeds the $Thresh_2^{busy}$, VDA will initiate handover for its data application as well. The basis for Policy II is that for voice and video streaming which require good QoS, congestion at the cheaper WLAN will necessitate that such applications be given top priority in a better network. Data on the other hand, can remain in the WLAN in a bid to reduce the cost of using WMAN. Intuitively, if the voice and video have been switched to WMAN, the WLAN congestion will ease somewhat resulting in better QoS for the data, unless the WLAN is really saturated. In such case, the data will also be switched to WMAN in Policy II. In both policies, handover back from WMAN to WLAN will depend on whether the WLAN network has become less busy through the intelligent monitoring of the prevailing network conditions.

The process for this handover is summarized in Algorithm 1. As more stations join the WLAN, the network becomes progressively loaded. Upon periodic update of L2 metrics, the framework notices the change as reflected by higher observed channel busy ratio. If the measured channel busy ratio crosses the second busy time threshold ($T_{tot(t)}^{busy} \geq Thresh_2^{busy}$), then VDA executes the selected policy for the handover of the applications. During handover preparation, VDA performs QoS mapping for appli-

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

cations to be properly handled in WMAN network. The voice and video streaming applications are adapted into the respective service flow characteristics as per the Traffic Specification (TSPEC). The terminal then establishes connections to WiMAX BS. After successful exchange of DSA-REQ/DSA-RSP (Dynamic Service Addition *request-response*) messages, VDA checks its connection table for existing sessions, then re-associates them with the WiMAX interface. The mapping strategy is shown in Table 5.2. Voice, video streaming and data application are handled using UGS, rtPS, and BE scheduling respectively in WMAN network. Note that in the scheme, we set $\alpha * Thresh_2^{busy}$ (with $\alpha = 0.9$) to be the threshold that triggers the handover decision. This is to ensure session seamlessness before the WLAN gets highly congested.

Algorithm 1 WLAN $\rightarrow$ WMAN handover

```

voice  $\leftarrow$  Wi-Fi
video  $\leftarrow$  Wi-Fi
data  $\leftarrow$  Wi-Fi
if  $T_{tot(t)}^{busy} \geq \alpha * Thresh_2^{busy}$  then
  if use Policy I then
    handover {voice, video streaming, data} to WMAN
  else if use Policy II then
    if (voice  $\leftarrow$  WiMAX) AND (video  $\leftarrow$  WiMAX) then
      handover {data} to WMAN
    else
      handover {voice, video} to WMAN
    end if
  end if
end if

```

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

5.3.2 WMAN $\rightarrow$ WLAN handover

The process for this handover is summarized in Algorithm 2. There is no difference between Policies I and II regarding this handover scenario. By default, user prefers WLAN service once it is available. However, the decision to move back from WMAN to WLAN based solely on service availability is less optimal. Intelligent terminal needs to consider a further criterion in order to enhance user perceived QoS and avoid unnecessary handover. In our scheme, we differentiate the handling mechanism for voice and video streaming from that of data transfer application. For voice and video streaming applications, handover back into WLAN network occurs only if channel busy time ratio is less than the first busy time threshold ($T_{tot(t)}^{busy} \leq Thresh_1^{busy}$) for a configurable period of time. The shift of channel busy time value (from $T_{tot(t)}^{busy} \geq Thresh_2^{busy}$ to $T_{tot(t)}^{busy} \leq Thresh_1^{busy}$) implies that significant gain, in terms of available bandwidth, can be provided to the application. The hysteresis window is applied to alleviate switching oscillations between the two interfaces due to transient network fluctuations, hence ensuring access network stability. Different approach for handover applies for data transfer application. Handover occurs if $Bw_Req_{loss(t)} \geq \alpha * Thresh_2^{bwreq}$ and $T_{tot(t)}^{busy} \leq Thresh_2^{busy}$. Such consideration for data application is required as this type of application is serviced using best effort connection without any guarantee of QoS. Therefore terminal shall find the best access network to accomodate such application.

Table 5.2: Mapping of QoS for different technologies

Example of Application	IEEE 802.11e	IEEE 802.16
Voice	AC_VO	UGS
Video streaming	AC_VI	rtPS
FTP	AC_BK	nrtPS
HTTP	AC_BE	BE

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

Algorithm 2 WMAN $\rightarrow$ WLAN handover

```

SET  $n \leftarrow 0$  ;  $window \leftarrow 3$ 

 $voice \leftarrow \text{WiMAX}$ 
 $video \leftarrow \text{WiMAX}$ 
 $data \leftarrow \text{WiMAX}$ 
if  $T_{tot(t)}^{busy} \leq Thresh_1^{busy}$  then
  if  $n == window$  then
    handover {voice, video streaming} to WLAN
  else
    increment  $n$ 
  end if
else if  $(BW\_REQ_{loss(t)} \geq \alpha * Thresh_2^{bwreq})$  AND  $(T_{tot(t)}^{busy} \leq Thresh_2^{busy})$  then
  if  $n == window$  then
    handover {data} to WLAN
  else
    increment  $n$ 
  end if
else
  remain in WMAN
end if

```

5.4 Simulation Results

We study the performance of our scheme using ns-2 network simulator [35], from which we observe metrics such as throughput and delay for voice, video streaming, and data type application. What is expected by integrating the framework with QoS support in the link layer is that the perceived quality, especially of voice and video applications, is preserved during the session duration. For applications in the QoS-enabled terminal, we use the following CBR traffic:

- voice application: 128 Kbps, with packet size of 320 bytes and interval rate of 0.02s
- video streaming application: 768 Kbps, with packet size of 960 bytes and interval rate of 0.01s
- data transfer application: 384 Kbps, with packet size of 960 bytes and interval rate of 0.02s

The channel busy ratio thresholds used in the simulation are as follows:

- $Thresh_1^{busy} = 0.376844$
- $Thresh_2^{busy} = 0.841756$

The bandwidth request loss ratio thresholds used in the simulation are as follows:

- $Thresh_1^{bwreq} = 0.266771$
- $Thresh_2^{bwreq} = 0.456171$

The frame duration and T16 timer (timer to indicate Bandwidth request expiration) in WiMAX is set to 5 ms and 100 ms respectively. Table 5.3 shows the configuration for ACs in the terminal according to [48]. The simulation is run until $t = 300s$. Applications in the QoS-enabled terminal start at $t = 15$. During the simulation run time, we continuously introduce Wi-Fi and WiMAX terminals sending 384

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

Table 5.3: AC setting in terminal

Access Category (AC)	AIFSN	CW _{Min}	CW _{Max}
AC_VO	2	7	15
AC_VI	2	15	31
AC_BE	3	31	1023
AC_BK	7	31	1023

kbps background traffic through the networks. At time $t = 20, 30, \dots, 70$ (at interval of 10 s), a Wi-Fi terminal is added, giving a total number of 7 background terminals in WLAN network. Meanwhile, at the same interval of 10 seconds, a WiMAX terminal is added at time $t = 25, 35, \dots$, and 105, giving a total number of 9 background terminals in the WMAN network. In the middle of the simulation at time $t = 180$, 5 Wi-Fi terminals leave the network, resulting in a decrease in the WLAN network load.

Fig. 5.3 to 5.5 show the results of using Policy I. As can be seen from Fig. 5.5, at $t = 71$, terminal handoffs all of its sessions from WLAN network to WMAN network when the channel busy ratio crosses $Thresh_2^{busy}$. At time $t = 89$, while the BW-Req loss ratio crosses $Thresh_2^{bwreq}$, terminal also notices that its channel busy ratio $< Thresh_2^{busy}$, hence it handovers its data application back to the WLAN network. Around time $t = 184$, upon detecting a gain in the available WLAN throughput (channel busy ratio $< Thresh_1^{busy}$) over the hysteresis window (the window is set to 3), the terminal handovers its voice and video streaming sessions back to the WLAN network. The throughput measurement shows clearly that QoS for the sessions can be maintained, both for WLAN-WMAN and WMAN-WLAN handover, due to a proper QoS mapping strategy. For the data transfer application, it can be seen that after several seconds of sending packets through the WMAN network, the terminal handovers the data back to the WLAN network to maintain the throughput.

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

For voice and video streaming sessions, delay determines the perception quality of applications at the receiver side. Figure 5.3 shows that delay for both voice and video streaming application can be maintained consistently accross the WLAN and WMAN networks.

Fig. 5.6 to 5.8 show the results when the terminal prioritizes its voice and video streaming application first for handover to WMAN network under Policy II. It can be seen that the data application still remains at WLAN network during the whole duration of the session. The delay measurement shows higher value during $t = 71$ to 184 as WLAN network gets more congested.

5.5 Summary

In this chapter we have adapted the AMIMF in Chapter 4 to incorporate efficient switching policies for effective multi-interface access so as to cater to the different user preferences in terms of cost versus quality as well as to cater to the application requirements. The results show that the terminal's ability to assess the access network performance based on L2 metrics measurement coupled with QoS provided by specific L2 technologies enhances user's perceived quality as session can be maintained during the user's movement.

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

Figure 5.3: Policy I: Delay in dual-interface terminal

Figure 5.4: Policy I: Throughput of network interfaces in dual-interface terminal

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

Figure 5.5: Policy I: Throughput of applications in dual-interface terminal

Figure 5.6: Policy II: Delay in dual-interface terminal

CHAPTER 5. QOS SUPPORT IN MULTIHOMED TERMINAL

Figure 5.7: Policy II: Throughput of network interfaces in dual-interface terminal

Figure 5.8: Policy II: Throughput of applications in dual-interface terminal

Chapter 6

Conclusion

6.1 Conclusion

Next generation wireless network has been envisioned to comprise heterogeneous wireless networks where different access technologies (Wi-Fi, WiMAX, UMTS, etc) will cooperate seamlessly. From the user's point of view, the network offers capability to experience a new way of communication. Given that the user's device is multihomed, user can preserve connectivity to any of the wireless access technologies which differ in terms of range, frequency, capacity, and QoS support. As the quality of experience is usually assessed subjectively based on user's perspective towards the network service, the selection of access network plays a vital role in determining the user's satisfaction. Notwithstanding the facts that radio resource management, which includes admission control and resource provisioning has been applied within the network, user can still choose policies which would choose the access network that best supports his demand in terms of application requirement, terminal capability, and other preferences (eg. cost, network priority, etc). On the other hand, in the absence of such user-centric policy framework in the network, user is still supported with best effort to benefit from whatever service the network has provided. Thus, regardless of whether or not certain QoS is provided through heterogenous networks, a smart device supporting

CHAPTER 6. CONCLUSION

user's application becomes a necessity.

In this thesis, our endeavor to boost the device's intelligence is achieved by composing a framework that enables cooperation of multiple interfaces in the device. The operation of our framework is mostly done at the end user device which requires minimum support from network entities. The framework is built upon the common Media Independent Handover standard, thus ensuring interoperability across multiple access technologies. To cater for mobility issue, the MIH based Multi-homed Mobility Management Framework (MHMM) employs MSIP. MHMM allows user to fully benefit from heterogeneous wireless networks and maintain the communication session when switching association from one access network to another. This inter-technology handover or vertical handover implies that the user selects the appropriate access network not only at the session initiation but also during the whole session to yield the greatest benefit.

We introduce one important module which is called VDA. This module, in cooperation with MIH function, performs significant tasks in the MHMM framework. Such tasks include application's connection state maintenance, vertical handover support by communicating with MSIP, and assignment of flow or packet to the network interface during transmission. The framework prevents session disruption by means of IP address mapping. By employing the mechanism, any IP changes is transparent to the application which is also useful especially for TCP-based application. During vertical handover, to anticipate longer delay incurred inherently by protocol residing at application layer, the framework makes use of MIH Link Event trigger which allows device to finish the handover execution before network disconnection and whenever applicable, maintains connectivities to more than one access network. Our framework exploits the MIH protocol to obtain dynamic link layer information from all interfaces. The information is further used as decision metrics in the transmission policy. We introduce adaptive load balancing policy that is applied specifically for application with idle and burst characteristic. Through this policy, the framework obtains

CHAPTER 6. CONCLUSION

L2 delay and frame loss metric and utilize SAW to derive the weight of associated access networks when selecting the interface to use at the beginning of new burst. Our result shows that the policy is effective in dealing with rapid change in access network condition.

MHMM is extended into an Adaptive Multiple-Interface Management Framework (AMIMF) where we equip the device with prior knowledge to help estimate the throughput available in an access network. This gleaned information can be obtained through non-intrusive scheme by utilizing a collection of trace history of previous transmissions. More metrics are extracted from underlying link layers such as L2 delay, channel busy time and BW-REQ loss ratio. The respective combination of observed parameters leads to us defining a set of access network classifications based on its performance in terms of available throughput. Moreover, we also show that the classification directly relates to the congestion level of the access network. To optimize multiple interface usage, several load balancing policies are proposed, such as RR, WRR, Adaptive distribution and Hybrid scheme. Based on the performance of these policies, we come to the suggestion that Wi-Fi interface should be deemed non-eligible for transmission when the WLAN network is highly congested as this condition results in more severe impact on application as compared to congested conditions in WMAN network. The result shows that awareness of the access network condition is non-trivial in the user's perceived quality enhancement, especially for application with long connection duration such as media streaming.

We have also integrated AMIMF with QoS link layer support. The QoS-integrated AMIMF facilitates different types of applications and ensures QoS consistency across wireless domains by mapping of application's requirement into its equivalent network service. The coupling of terminal's ability to assess the access network performance based on L2 metrics measurement and QoS provided by specific L2 technologies results in an improvement to user perceived quality.

CHAPTER 6. CONCLUSION

6.2 Recommendation for Future Work

While our framework currently supports L2 QoS provisioning, the work can be extended to provide solutions for end-to-end QoS for applications. There exist several works in the literature which address end-to-end QoS issue. Jung et al. [49] propose using SIP to handle end-to-end QoS setup for each multimedia session. When vertical handover occurs, QoS of application may not be maintained due to the different available throughput in the new network. The scheme integrates SIP and MIH for the purpose of traffic flow adaption and QoS reconfiguration after vertical handover. The scheme relies on SIP Re-Invite message to modify the current QoS configuration. The scheme supports segmented network model, where the QoS session setup is done separately for local and remote access network.

An insight of extending the operation of MIH for QoS reservation is presented in [50]. MIH_Handover_Prepare.request primitive is used not only to query available resource but also to include QoS specific parameters in the preparation phase to allow reservation of QoS in specific link layer. The authors propose a new MIH command set, MIH_Resource_Prepare.request/response for the session setup. In the QoS session setup, reservation on the network is never triggered by the terminal for security reason, but it has to be initiated by a trusted network entity. The implication would be for terminal to send the QoS reservation request to the core network.

Baek et al. [51] propose a QoS architecture for the next generation wireless network based on DiffServ model which is achieved through two components, Access Network QoS manager (AQM) and Inter-AQM (IAQM). AQM manages local QoS in each access network while IAQM controls QoS between heterogeneous access network. Each access network is connected through an IP core network, where path between terminal and CN will traverse the core network and wireless access network. Thus QoS mapping mechanism is performed between access network with its specific QoS mechanism and Core network with the DiffServ mechanism, thus ensuring end-to-end QoS for traffic flow.

CHAPTER 6. CONCLUSION

In [52], a framework to provide an end-to-end QoS solution for terminal in a mobility scenario is proposed. The framework decouples the application signaling from the network signaling. Implementing end-to-end SIP signaling, the framework enables the applications to adapt dynamically to the terminal's current QoS connection. Meanwhile, for the network QoS signaling and reservation, a centralized QoS architecture is adopted, where there exists a domain resource manager (DRM) responsible for QoS management in each domain. Both DiffServ and IntServ can be applied in the core network.

In this thesis, our framework addresses outbound packet transmission over multiple interfaces where multihomed device constitutes the sender. However, we do not consider the case where the multihomed device functions as the receiver. Therefore we propose further work on load balancing over multiple interfaces for the device to govern the inbound packet reception. For this purpose, a proxy server is required to be deployed in the network which performs demultiplexing of flow associated with the user. Link layer parameters from our framework can be used as QoS feedback for the proxy which is realized through communication between MIHF residing in the user device and the proxy server.

Author's Publications

1. Widyo Cahyono Andi, Yang Xia, Chai Kiat Yeo, Bu Sung Lee, MIH-based SIP Mobility Management Scheme in Heterogeneous Wireless Networks, in Proc. The Second International Conference on Mobile Ubiquitous Computing, System, Services and Technologies (UBICOMM), Valencia, Spain, 29 Sep - 4 Oct 2008.
2. Widyo Cahyono Andi, Chai Kiat Yeo, Bu Sung Lee, Environment-Aware QoS Framework for Multi-Interface Terminal, accepted for publication in Journal: Computer Communications, by Elsevier Science, January 2010.
3. Widyo Cahyono Andi, Yang Xia, Chai Kiat Yeo, Bu Sung Lee, An Interface Management Framework for Multihomed Terminal, in Proc. 7th International Conference on Information, Communications and Signal Processing (ICICS 09), Macau, 7 - 10 Dec 2009.
4. Widyo Cahyono Andi, Yang Xia, Chai Kiat Yeo, Bu Sung Lee, An Adaptive Terminal-based Multiple-interface Management Framework, submitted for review to Journal: Wireless Communications and Mobile Computing, by Wiley, Jul 2009.

Bibliography

- [1] T. Ernst, N. Montavont, R. Wakikawa, E. Paik, C. Ng, K. Kuladinithi, and T. Noel. Goals and Benefits of Multihoming. draft-multihoming-generic-goals-and-benefits-02). October 2005.
- [2] AI.F. Akyildiz, J. McNair, J.S.M. Ho, H. Uzunalioglu, W. Wang,. Mobility management in next-generation wireless systems. *IEEE Proceedings*, 87(8):1347–1384, Aug 1999.
- [3] M. Li, K. Sandrasegaran, and T. Tung. A multi-interface proposal for IEEE 802.21 media independent handover. In *Proceedings of International Conference on the Management of Mobile Business*, Toronto, Canada, July 2007.
- [4] Draft IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services. *IEEE LAN/MAN Draft IEEE P802.21/D00.05*, January 2006.
- [5] T. You, and S. Lee. An approach on making use multiple interface of mobile node simultaneously. In *9th International Conference on Advanced Communication Technology (ICACT 2007)*, Phoenix Pk, South Korea, Feb 12-14, 2007.
- [6] F. Jover, G. Reid, and X. Jover. A faster Handover Mechanism using SIP. In *Proceedings of The16th IST Mobile And Wireless Communications, Vols 1-3*, pages 444–448. Budapest, Hungary, Jul 01-05, 2007.

BIBLIOGRAPHY

- [7] F. Andre, J. M. Bonnin, B. Deniaud, K. Guillouard, N. Montavont, T.Noel, and L. Suciu. Optimized support for multiple wireless interfaces within an IPv6 end-terminal. In *the Smart Objects Conf.*, Grenoble, France, May 15-17, 2007.
- [8] N. Fikouras, K. Kuladinithi, C. Gorg, C. Bormann, and A. Timm-Giel. Multiple Access Interface Management and Flow Mobility. In *Proc. of the 13th IST Mobile and Wireless Communications Summit 2004*, Lyon, France, 2004.
- [9] G. Lampropoulos, A. K. Salkintzis, and N. Passas. Media-independent handover for seamless service provision in heterogeneous networks. *IEEE Communications Magazine*, vol. 46, no. 1, pp.64-71, January 2008.
- [10] L. Eastwood, S. Migaldi, Q. Xie, and V. Gupta, Mobility using IEEE 802.21 in a heterogeneous IEEE 802.16/802.11-based, IMT-advanced (4G) network. *IEEE Wireless Communications*, 15(2):26–34, April 2008.
- [11] W. Lee, M.-S. Kang, and M.-S. Lim. Implementation of 802.21 for seamless handover across heterogeneous networks. In *Managing Next Generation Networks and Services*, volume 4773 of *Lecture Notes in Computer Science*, pages 326-333, 2007. Proceedings 10th Asia-Pacific Network Operations and Management Symposium, Sapporo, Japan, Oct 10-12, 2007.
- [12] B.-S. Lee, T. M. Lim, C. K. Yeo, and Q. V. Le. A mobility scheme for personal and terminal mobility. In *Proc. IEEE ICC*, pages 1815-1820, Glasgow, Scotland, June 2007.
- [13] T. M. Lim, C. K. Yeo, B.-S. Lee, and Q. V. Le. TMSP: Terminal Mobility Support Protocol. In *IEEE Transactions on Mobile Computing*, 8(6), pages 849-863, June 2009.
- [14] A. Abd El Al, T. N. Saadawi, and M. J. Lee, LS-SCTP: A bandwidth aggregation technique for stream control transmission protocol. *Computer Communications*, 27(10):pp.1012–1024, 2004.

BIBLIOGRAPHY

- [15] L. Magalhaes and R. Kravets. Transport level mechanisms for bandwidth aggregation on mobile hosts. In *Ninth International Conference on Network Protocols (ICNP 2001)*., pages 165–171, Nov. 2001.
- [16] H.-Y. Hsieh and R. Sivakumar. pTCP: An end-to-end transport layer protocol for striped connections. In *Proceedings. 10th IEEE International Conference on Network Protocols (ICNP 2002)*, pages 24–33, Nov. 2002.
- [17] Y. H. Lin and S.-L. Tsao. Dynamic bandwidth aggregation for a mobile device with multiple interfaces. In *The 8th International Symposium on Communications (ISCOM2005)*, Taiwan, 2005.
- [18] H. Huang, J. Cai, A. Kassler, and C. Fu. Load-sharing in wireless multi-homed systems. In *IEEE International Conference on Communications, 2005 (ICC 2005)*, volume 5, pages 3489–3493 Vol. 5, May 2005.
- [19] K. Manousakis and D. Famolari. Intelicon: A framework for the simultaneous utilization of multiple interfaces and its application on tcp. In *International Wireless Communications and Mobile Computing Conference, IWCMC '08.*, pages 976–981, Aug. 2008.
- [20] A. Snoeren. Adaptive inverse multiplexing for wide-area wireless networks. In *GLOBECOM'99: Seamless Interconnection for Universal Services, Vol 1-5*, pages 1665–1672, 1999. IEEE Global Communications Conference (GLOBECOM 99), Rio Janeiro, Brazil, Dec 05-09, 1999.
- [21] Mosawi TA, Wisely D, and Aghvami H. A novel micro mobility solution based on media independent handover and sip. In *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th*, pages 1–5, Sept. 2006.
- [22] W. Ying, Y. Jun, Z. Yun, L. Gen, and Z. Ping. Vertical handover decision in an enhanced media independent handover framework. In *Wireless Communications and Networking Conference, 2008. IEEE WCNC 2008.*, pages 2693–2698, 2008.

BIBLIOGRAPHY

- [23] R. Tawil, G. Pujolle, and J. Demerjian. Distributed handoff decision scheme using mih function for the fourth generation wireless networks. In *3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008.*, pages 1–6, April 2008.
- [24] H. Haffajee and H. A. Chan. Low-cost qos-enabled wireless network with interworked wlan and wimax. In *Proc. of the First IEEE Intl. Conf. on Wireless Broadband and Ultra Wideband Communications (AusWireless 2006)*.
- [25] Moby dick: Mobility and differentiated services in a future ip network. eu project ist-2000-25394, january 2003. <http://www.ist-mobydick.org>.
- [26] T. Ali-Yahiya, K. Sethom, and G. Pujolle. Seamless continuity of service across WLAN and WMAN networks: Challenges and performance evaluation. In *Standards, Architectures and Managements of Broadband Convergence Networks*, pages 100–111, 2007. 2nd IEEE/IFIP International Workshop on Broadband Convergence Networks, Munich, Germany, Mar 21, 2007.
- [27] C.-Y. Ho and H.-L. Chao. *An implementation of QoS Framework for Heterogeneous Networks*. In *Emerging Directions in Embedded and Ubiquitous Computing* 2007.
- [28] C. Perkins. IP Mobility Support. *RFC2002*, October 1996.
- [29] D. Johnson, C. Perkins and J. Arkko. Mobility Support in IPv6. *RFC3775*, June 2004.
- [30] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. SIP: Session Initiation Protocol. *RFC3261*, June 2002.

BIBLIOGRAPHY

- [31] H. Schulzrinne, and E. Wedlund. Application-layer mobility using SIP. *ACM SIGMOBILE Mobile Comp. and Commun. Rev.*, vol. 4, no. 3, pp.47-57, July 2000.
- [32] Seamless and Secure Mobility tool suite. <http://w3.antd.nist.gov/seamlessandsecure/>.
- [33] 3GPP TS 23.234 v6.3.0. 3GPP System to WLAN Interworking; System Description (Release 6). Mar. 2004.
- [34] L. Shi, B. Liu, C. Sun, Z. Yin, L. Bhuyan, and H. J. Chao. Flow-slice: a novel load-balancing scheme for multi-path switching systems. In *ANCS '07: Proceedings of the 3rd ACM/IEEE Symposium on Architecture for networking and communications systems*, pages 45–46, New York, NY, USA, 2007.
- [35] NS-2. The Network Simulator. <http://www.isi.edu/nsnam/ns/>.
- [36] Al Morton. Packet Delay Variation Applicability Statement. *draft-morton-ippm-delay-var-as-01.txt*, 29-Nov-06.
- [37] Y. Bejerano, S.-J. Han, and L. Li. Fairness and load balancing in wireless lans using association control. *IEEE/ACM Transactions on Networking*, 15(3):560–573, June 2007.
- [38] A. P. Jardosh, K. N. Ramachandran, K. C. Almeroth, and E. M. Belding-Royer. Understanding congestion in ieee 802.11b wireless networks. In *IMC '05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, pages 25–25, Berkeley, CA, USA, 2005. USENIX Association.
- [39] B. Melander, M. Bjorkman, and P. Gunningberg. A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In *Global Telecommunications Conference. IEEE GLOBECOM '00.*, vol.1, pages 415–420, 2000.

BIBLIOGRAPHY

- [40] S. Banerjee and A. K. Agrawala. Estimating available capacity of a network connection. In *Proceedings. IEEE International Conference on Networks, 2000. (ICON 2000).*, pages 131–138, 2000.
- [41] J. He, K. Guild, K. Yang, and H.-H. Chen. Modeling contention based bandwidth request scheme for ieee 802.16 networks. *IEEE Communications Letters*, 11(8):689–700, August 2007.
- [42] Widyo Cahyono Andi, Yang Xia, Chai Kiat Yeo, and Bu Sung Lee. MIH Based SIP Mobility Management Scheme in Heterogeneous Wireless Networks. In *UBICOMM '08: Proceedings of The Second International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, pages 7–12, Washington, DC, USA, 2008. IEEE Computer Society.
- [43] J.-S. Wu, S.-F. Yang, B.-J. Hwang. A terminal-controlled vertical handover decision scheme in ieee 802.21-enabled heterogeneous wireless networks. *International Journal of Communication Systems*, 9999(9999):n/a, June 2009.
- [44] IEEE Std. 802.11-1999. Local and metropolitan area networks specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications. *IEEE*, 1999.
- [45] Part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications: Amendment 7: Medium access control (mac) quality of service (qos) enhancements. *IEEE Unapproved Draft, Rev. D12.0*, 2004.
- [46] Intel White paper. Providing QoS in WLANs. http://intel.com/network/connectivity/resources/doc_library/white_papers/303762.htm.
- [47] IEEE Std 802.16-2004. IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. 2004.

BIBLIOGRAPHY

- [48] A. Andreadis, G. Benelli, and R. Zambon. Evaluation of qos support for multimedia traffics in ieee 802.11e. In *Proceedings of International Conference on Software in Telecommunications and Computer Networks, 2006. SoftCOM 2006.*, pages 91–95, 29 2006-Oct. 1 2006.
- [49] Y.-C. Jung, B.-K. Kim, and Y.-T. Kim. Sip based end-to-end qos negotiation scheme for mih. In *2nd IEEE/IFIP International Workshop on Broadband Convergence Networks, 2007. BcN '07.*, pages 1–12, May 2007.
- [50] Miguel Almeida, Daniel Corujo, Susana Sargento, Vitor Jesus, Rui Aguiar. An end-to-end QoS framework for 4G mobile heterogeneous environments. In *Open-Net Workshop*, pages 1–12, March 27-29, 2007, Diegem, Belgium.
- [51] S.-K. Baek, J.-U. Kim. A QoS management architecture in heterogeneous wireless networks environment *Wireless World Research Forum Meeting 21*, Stockholm, Sweden, 13-15 October 2008.
- [52] J. W. Tantra, M. N. Son, T. M. Lim, B.-S. Lee, and C. K. Yeo. Qos in terminal-assisted mobility. In *6th International Conference on Information, Communications and Signal Processing, 2007*, pages 1–5, Dec. 2007.