Universidade de Aveiro 2007

Departamento de Electrónica, Telecomunicações e Informática

Daniel Nunes Corujo

IEEE 802.21 in Heterogeneous Handover Environments

Universidade de Aveiro Departamento de Electrónica, Telecomunicações e Informática
2007

Daniel Nunes Corujo

IEEE 802.21 in Heterogeneous Handover Environments

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requesitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica do Prof. Dr. Rui Aguiar, Professor auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.

o júri / the jury

presidente / president

Prof. Dr. Valery Sklyarov
Professor Catedrático do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro

Prof. Dra. Marilia Pascoal Curado
Professora Auxiliar da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Prof. Dr. Rui L. Aguiar
Professor Associado do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro

agradecimentos / acknowledgements

Ao Professor Rui Aguiar pelo apoio, partilha de conhecimento e espírito crítico que foram fulcrais na elaboração desta dissertação. E, muito principalmente, pela sua docência e visão singular do mundo das redes de telecomunicações, que me inspiraram em todo o meu percurso académico.

À Professora Susana Sargento pela sua disponibilidade e apoio sobre todos os assuntos que surgissem no desenvolvimento desta dissertação.

Ao Instituto de Telecomunicações por me ter fornecido todas as condições e apoio necessários ao desenvolvimento desta dissertação. Em especial, aos colegas e amigos do grupo de redes heterogéneas HNG (Heterogeneous Networking Group) por todo o seu interesse e apoio.

Ao Telemaco Melia, Albert Vidal e Antonio De La Oliva que contribuiram para a minha formação sobre o IEEE 802.21.

À minha namorada, Lara, e à minha familia por todo o apoio e motivação.

Palavras Chave

IEEE 802.21, Media Independent Handovers, NIHO, MIHO, Terminais Multimodo, Arquitecturas de Redes 4G

Resumo

O desenvolvimento das capacidades tecnológicas dos terminais móveis, e das infra-estruturas que os suportam, potenciam novos cenários onde estes dispositivos munidos com interfaces de diferentes tecnologias vagueiam entre diferentes ambientes de conectividade. É assim necessário providenciar meios que facilitem a gestão de mobilidade, permitindo ao terminal ligar-se da melhor forma (i.e., optando pela melhor tecnologia) em qualquer altura.

A norma IEEE 802.21 está a ser desenvolvida pelo *Institute of Electrical and Electronics Engineers* (IEEE) com o intuito de providenciar mecanismos e serviços que facilitem e optimizem *handovers* de forma independente da tecnologia. A norma 802.21 especifica assim um conjunto de mecanismos que potenciarão cenários como o descrito acima, tendo em conta a motivação e requerimentos apresentados por arquitecturas de redes futuras, como as redes de quarta geração (4G).

Esta dissertação apresenta uma análise extensiva da norma IEEE 802.21, introduzindo um conjunto de simulações desenvolvidas para estudar o impacto da utilização de mecanismos 802.21 em *handovers* controlados por rede, numa rede de acesso mista composta por tecnologias 802.11 e 3G. Os resultados obtidos permitiram verificar a aplicabilidade destes conceitos a ambientes de próxima geração, motivando também uma descrição do desenho de integração de mecanismos 802.21 a arquitecturas de redes de quarta geração.

Keywords

IEEE 802.21, Media Independent Handovers, NIHO, MIHO, Multimode Terminals, 4G Network Architectures

Abstract

The development of the technological capabilities of mobile terminals, and the infra-structures that support them, enable new scenarios where these devices using different technology interfaces roam in different connectivity environments. This creates a need for providing the means that facilitate mobility management, allowing the terminal to connect in the best way possible (i.e., by choosing the best technology) at any time.

The IEEE 802.21 standard is being developed by the Institute of Electrical and Electronics Engineers (IEEE) to provide mechanisms and services supporting Media Independent Handovers. The 802.21 standard specifies a set of mechanisms that enable scenarios like the one described above, considering the motivation and requirements presented by future network architectures, such as the ones from fourth generation networks (4G).

This thesis presents an extensive analysis of the IEEE 802.21 standard, introducing a set of simulations developed for studying the impact of using 802.21 mechanisms in network controlled handovers, in a mixed access network composed of 802.11 and 3G technologies. The obtained results allow the verification of the applicability of these concepts into next generation environments, also motivating the description of the design for integration of 802.21 mechanisms to fourth generation networks.

Contents

List of Figures

List of Tables

Acronyms

Acronym	Description
2G	Second Generation Networks
3G	Third Generation Networks
3GPP	Third Generation Partnership Project
4G	Fourth Generation Networks
A4C	Authentication, Authorization, Accounting, Auditing and Charging
AAA	Authentication, Authorization and Accounting
AAL	ATM Adaptation Layer
AP	Access Point
API	Application Programming Interface
AR	Access Router
ASN.1	Abstract Syntax Notation number One
AN	Access Network
ATM	Asynchronous Transfer Mode
BA	Binding Acknowledgement
BAsh	Bourne-again Shell
BE	Best Effort
BS	Base Station
BSS	Basic Service Set
BU	Binding Update
CDMA	Code Division Multiple Access
CID	Connection Identifier
CIP	Cellular IP
CN	Core Network
CoA	Care-of Address
CPS	Common Part Sublayer
CS	Convergence Sublayer
DAD	Duplicate Address Detection
DAIDALOS	Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimised personal Services
DDS	Distribution System
DHCP	Dynamic Host Configuration Protocol
DSM	Distribution System Medium
EDGE	Enhanced Data rates for Global Evolution
ESS	Extended Service Set
FDD	Frequency Division Duplex
GERAN	GSM-EDGE Radio Access Network
GGSN	GPRS Support Node

Continued on next page

continued from previous page

Acronym	Description
GLR	Gateway Location Register
GMD	Global Mobility Domain
GMP	Global Mobility Protocol
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HA	Home Agent
HIP	Host Identity Protocol
HMIPv6	Hierarchical MIPv6
HoA	Home Address
IBSS	Independent Basic Service Set
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISP	Internet Service Provider
IST	Information Society
LAN	Local Area Network
LLC	Logical Link Control
LMA	Localized Mobility Anchor
LMD	Local Mobility Domain
LMM	Localized Mobility Management
LMP	Local Mobility Protocol
LoS	Line of Sight
MAC	Medium Access Control
MAG	Mobility Access Gateway
MAN	Metropolitan Area Network
MARQS	Mobility Management, AAA, Resource Management, QoS and Security
MDNS	Multicast Domain Name System
MIB	Management Information Base
MICS	Media Independent Command Service
MIES	Media Independent Event Service
MIH	Media Independent Handovers
MIHF	Media Independent Handovers Function
MIHO	Mobile Initiated Handovers
MIIS	Media Independent Information Service
MING	Media Independent Neighborhood Graph
MIP	Internet Protocol Mobility Support
MIPv6	Internet Protocol Mobility Support for IPv6
MLME	MAC Layer Management Entity
MME	Mobility Management Entity
MN	Mobile Node
MPLS	Multi-Protocol Label Switching
MSC	Mobile Switching Center
MSDU	MAC Service Data Unit
NAI	Network Address Identifier
NAP	Network Access Point
NAT	Network Address Translation

Continued on next page

continued from previous page

Acronym	Description
NCMS	Network Control Management System
NetLMM	Network Localized Mobility Management
NIHO	Network Initiated Handovers
NMS	Network Management System
NPDB	Number Portability DataBase
nrtPS	non real-time Polling Service
NSIS	Next Steps in Signaling
OSI	Open Systems Interconnection
PBU	Proxy Binding Update
pCoA	Proxy Care-of Address
PDU	Protocol Data Unit
PDP	Protocol Data Packet
pHoA	Proxy Home Address
PHY	Physical Layer
PLME	Physical Layer Management Entity
PLMN	Public Land Mobile Network
PMA	Proxy Mobile Agent
PMIPv6	Proxy Mobile IPv6
PoA	Point of Attachment
PoS	Point of Service
PSTN	Public Switched Telephone Network
QAR	QoS-Authorization Request
QAA	QoS-Authorization Answer
QoS	Quality of Service
QoSC	QoS Client
QoSM	QoS Manager
RAL	Radio Access Layer
RNC	Radio Network Controller
RSS	Radio Signal Strength
RSSI	Radio Signal Strength Indication
rtPS	real-time Polling Service
RTT	Round-Trip Time
SAP	Service Access Point
SGSN	Service GPRS Support Node
SHIM6	Site Multihoming by IPv6 Intermediation
SIP	Session Initiation Protocol
SLP	Service Location Protocol
SMS	Short Message Service
SNMP	Simple Network Management Protocol
SS	Subscriber Station
SSID	Service Set Identifier
STA	Station
TCP	Transport Control Protocol
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TLV	Type Length Value
UDP	User Datagram Protocol
UGS	Unsolicited Grant Service

Continued on next page

continued from previous page

Acronym	Description
UMA	Unlicensed Mobile Access
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
VHE	Virtual Home Environment
VLR	Visitor Location Register
WCDMA	Wide Band Code Division Multiple Access
WEP	Wired Equivalent Privacy
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WM	Wireless Medium
XML	eXtensive Markup Language
ZQoSB	Zone QoS Broker

Chapter 1

Introduction

"The Right Tool for the job" is a commonly employed phrase in engineering throughout many different areas. It involves more than just a selection of what is available to execute a certain task. It combines the effort of assessing a number of issues and requirements that lead to the development of a specific mean for delivering the optimal execution of whatever task or mechanism needs to be done.

In this chapter the reader will be briefly introduced to the subjects that have motivated the work done on this thesis, providing an environment surrounding the task (or "job", as the phrase mentions), which will present a series of challenges and requirements for modern-day telecommunication networks concepts. The "tool", IEEE 802.21 [1] is presented and thoroughly studied in this work, considering not only the presented requirements but its application into european-wide design architectures and standardization mechanisms will also be presented.

1.1 Environment

The fast pace of modern life enforces technology to evolve more and more, granting empowering means for the common man that enable him to perform tasks in a better, faster and more satisfying way. Central to that, modern telecommunications allow users to be reachable almost everywhere, able to exchange important data across the world. However, in the future to be, that just isn't enough. New paradigms arise, where users desire accessibility to the Internet in an ubiquitous way, able to communicate anytime, anywhere and in anyway mean possible. This not only empowers users but also presents business opportunities for operators as well as potentiates new services from service providers.

In order to support this soon-to-be "ideal" a background research and conception effort is ever present. Mobile communication technologies have evolved into cellular and wireless networks, offering a plethora of devices that co-exist, each with it's own intrinsic and interesting capabilities and characteristics. However, the next big challenge is the combination of all these technologies into a single mean of access provision, allowing user devices equipped with several interfaces to be connected through the best access possible, with the best throughput and the least cost.

This heterogeneous environment, however, requires technological support in the form of an operator (or several operators) network(s) architecture, able to cope with user mobility and demand. More over, concepts such as Quality of Service have to be considered, in order to provide a gratifying experience to the user.

These considerations represent the motivation for studying an *enabler* that aims to smooth the effort that brings these concepts together. That *enabler* is the IEEE 802.21 standard, which is the major focus of this thesis. The study of this subject was further enhanced with the application of 802.21 mechanisms to the design of systems that supported mobility scenarios and 4G networks.

1.2 IEEE802.21

The IEEE802.21 standard, or Media Independent Handovers, presents the means and mechanisms to support technology abstraction to higher layers, which typically are in charge of control and decision-making, affecting areas such as mobility, admission control and load balancing, that are performed in the network or directly at terminals. The standard aims at providing information with the objective of enhancing handover procedures. In several chapters of this thesis, this standard is explored and applied to a simulator and insighting results are obtained.

1.3 Publications

The work developed under the subject of this thesis was published in several conferences. Results on a simulator implementation of 802.21 Network Controlled Handover Mechanisms in Heterogeneous Networks was published in [2] at the *IEEE Wireless Communications and Networking Conference 2007* and presented in Hong Kong this March. The conceptual application of 802.21 mechanisms to DAIDALOS[1], a Information Society (IST) *6th Framework Program* (FP6) integrated project, also resulted in publications at the *OpenNET Workshop* in Brussels [3], the *IEEE Symposium on Computers and Communications (ISCC'07)* [4] in Aveiro and the *16th Mobile and Wireless Communications Summit (MobileSummit 2007)* in Budapest. Also, pending acceptance, a paper was submitted to the *Elsevier Computer Networks Journal* [5].

Lastly, an Internet Draft [6] was also submitted to the *Internet Engineering Task Force (IETF)*, more concretely the *Mobility for IP: Performance, Signaling and Handoff Optimization (MIPSHOP)* Work-group, which was presented in Prague at the *68th IETF Meeting*, this March. A new version of this draft is currently under development to be presented to the *Network-based Localized Mobility Management* Work-group.

1.4 Document Outline

The remainder of this document is organized as follows: Chapter 2 introduces the reader to the concepts that motivated the study of the 802.21 standard, where an overview is provided in specific technologies and mobility mechanisms and how their integration into Quality of Service enabled-architectures is possible. Chapter 3 presents the IEEE 802.21 standard with it's concepts, services and mechanisms. In chapter 4 an existing simulator is analyzed, where a mobile terminal is equipped with 802.21 technology enabling it to execute optimized handovers between 3G and WLAN networks. Results are also studied. In chapter 5 the core of the work is presented, where the simulator presented in the previous chapter was extended with new code to support network controlled handovers, presenting new simulation results. Chapter 6 presents an overview of the application of 802.21 mechanisms to 4G network architectures. In chapter 7 final considerations conclude the work done in this thesis. Lastly, the Appendix presents some descriptions of the OMNeT++ Simulator in section A and implementation details in section B.

[1] http://www.ist-daidalos.org/

Chapter 2

Supporting a Mobile World

With the early development of wireless and cellular networks a new paradigm in communication became available for users that were mostly concerned with the public standard telephony network. Users started to be mobile and since then wireless communications gained more and more importance.

A plethora of new wireless technologies enables new scenarios where users are no longer "fixed", but always on the move, carrying with them their on-going connections and calls. These wireless technologies, however, have limitations: some can present low coverage areas or low throughput, and many do not take into consideration Quality of Service, to maintain a good user experience. It is verifiable that a single technology is not able to economically accommodate all the demands from users and applications. Thus a combination of several technologies is required, giving birth to new devices supporting multiple technologies and with the ability to select the most appropriate one at each time. However, increased and multiple link layer capability is not enough. A common ground had to be found which enabled applications and entities residing "over" the technology to communicate in an uniform way. This is achieved through usage of the Internet Protocol (IP) as the convergence layer for data and voice communication. Technology specific mobility protocols have obvious limitations when different kinds of technologies are at stake. Thus, although multiple proposals have shown increased benefits of spectrum and network usage [7], [8], [9], they cannot be directly used in a multiple technology environment. Layer 3 protocols are required for this environment, and in particular IP-based protocols. As such, considering the mobility environments pointed out earlier, mobility over the IP protocol, or Mobile IP, has to be considered, as well as it's implications on new network architectures. The combination of these concepts have been the source for the conceptualization of 4G architectures, enabling a common approach that supports several technologies, mobility and Quality of Service.

In this chapter an exploration of these concepts is proposed, empowering the reader with the requirements and facilities that enable a future heterogeneous environment.

2.1 Wireless Technologies

In the recent years, wireless Internet access via e.g. Wireless LAN has become widely spread both in private/corporate or public environments, making Internet access both mobile and ubiquitous. The merging of the Internet world together with the wireless communication world, combined with the availability of mobile devices supporting multiple wireless technologies (such as Wireless LAN and 2G/3G) created new business opportunities for mobile operators. In this new world, users can access services anywhere and anytime, always being best connected while enjoying a great variety of applications. In this section a brief overview of three relevant wireless technologies is presented.

2.1.1 IEEE 802.11

This section briefly highlights the architecture and common services of the IEEE 802.11 standard.

An 802.11 LAN is based on a cellular architecture where the system is subdivided into cells, where each cell (called Basic Service Set or BSS, in the 802.11 nomenclature) is controlled by a Base Station (called Access Point, or in short AP). The BSS consists of a coverage area within which the member stations (STA) of the BSS may remain in communication. If a station moves out of its BSS, it can no longer directly communicate with other members of the BSS. The independent BSS (IBSS) is the most basic type of IEEE 802.11 LAN.

To become a member of an infrastructure BSS, a station shall become "associated". These associations are dynamic and involve the use of a distribution system service (DSS).

A minimum IEEE 802.11 LAN may consist of only two stations. Because this type of IEEE 802.11 LAN is often formed without pre-planning, this type of operation is often referred to as an ad hoc network.

Even if a wireless LAN may be formed by a single cell, with a single Access Point, most installations will be formed by several cells, where the Access Points are connected through some kind of backbone (called Distribution System or DS), typically Ethernet, or in some cases wireless itself. IEEE 802.11 logically separates the wireless medium (WM) from the distribution system medium (DSM). The DS enables mobile device support by providing the logical services necessary to handle address to destination mapping and seamless integration of multiple BSSs. An access point (AP) is a STA that provides access to the DS by providing DS services in addition to acting as a STA. Data is transferred between a BSS and the DS via an AP. Note that all APs are also STAs, thus they are addressable entities. The addresses used by an AP for communication on the WM and on the DSM are not necessarily the same.

The whole interconnected Wireless LAN, including the different cells, their respective Access Points and the Distribution System, is seen to the upper layers of the OSI model, as a single 802 network, and is called in the standard [10] as an Extended Service Set (ESS). The key concept is that the ESS network appears the same to an LLC layer as an IBSS network. Stations within an ESS may communicate and mobile stations may move from one BSS to another (within the same ESS) transparently to the Logical Link Control (LLC).

There are two categories of IEEE 802.11 services: the station service (SS) and the distribution system service (DSS). Both categories of service are used by the IEEE 802.11 MAC sublayer. This set of services is divided into two groups: those that are part of every STA, and those that are part of a DS. Each of the services is supported by one or more MAC frame types. Some of the services are supported by MAC management messages and some by MAC data messages.

To integrate the IEEE 802.11 architecture with a traditional wired LAN, a final logical architectural component is introduced, the bridge. A bridge is the logical point at which Media Access Control Service Data Units (MSDUs) from an integrated non-IEEE LAN enter the IEEE 802.11 DS.

2.1.1.1 IEEE802.11 protocol stack layers

The 802.11 protocol stack layers have this form:

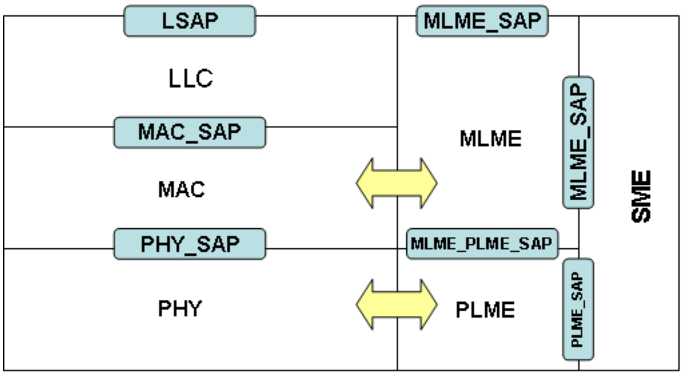

Figure 2.1: 802.11 layers

The 802.11 physical layer (PHY) is the interface between the MAC and the wireless media where frames are transmitted and received. The PHY provides three functions. First, the PHY provides an interface to exchange frames with the upper MAC layer for transmission and reception of data. Secondly, the PHY uses signal carrier and spread spectrum modulation to transmit data frames over the media. Thirdly, the PHY provides a carrier sense indication back to the MAC to verify activity on the media.

The 802.11 MAC layer provides functionality to allow reliable data delivery for the upper layers over the wireless PHY media. The data delivery itself is based on an asynchronous, best-effort, connectionless delivery of MAC layer data. There is no guarantee that the frames will be delivered successfully. It supports mechanisms such as scanning, authentication, association, WEP and fragmentation.

The LLC provides end-to-end link control over an 802.11-based Wireless LAN. It supports services such as Unacknowledged Connectionless Service, Connection-Oriented Service and Acknowledged Connectionless Service.

The PLME and MLME are the Physical Layer Management Entity and MAC Layer Management Entity, respectively. This entities supply access to information about administration in a MIB (Management Information Base).

The SAPs (Service Access Points) provide the interfaces between the different layers. More information can be obtained in [10].

2.1.1.2 Station service (SS)

The SS is present in every IEEE 802.11 station (including APs, as APs include station functionality).

- **Authentication** This service is used by all stations to establish their identity to other stations with which they will communicate. This is true for both ESS and IBSS networks. If a mutually acceptable level of authentication has not been established between two stations, an association shall not be established.

- **De-authentication** The de-authentication service is invoked whenever an existing authentication is to be terminated.

- **Privacy** IEEE 802.11 specifies an optional privacy algorithm, WEP, that is designed to satisfy the goal of wired LAN "equivalent" privacy. The algorithm is not designed for ultimate security but rather to be "at least as secure as a wire". To bring the functionality of the wireless LAN up to the level implicit in wired LAN design, IEEE 802.11 provides the ability to encrypt the contents of messages. Recent amendments to the standard have improved security significantly with the introduction of WPA2 in IEEE 802.11i.

- **MSDU Delivery** Delivery of MAC Service Data Units

2.1.1.3 Distribution System Service (DSS)

Services are used to cross media and address space logical boundaries. The DSSs are provided by the DS. They are accessed via a STA that also provides DSSs. A STA that is providing access to DSS is an AP.

- **Association** Before a STA is allowed to send a data message via an AP, it shall first become associated with the AP. The act of becoming associated invokes the association service, which provides the STA to AP mapping to the DS. The DS uses this information to accomplish its message distribution service.

- **Disassociation** The disassociation service is invoked whenever an existing association is to be terminated. In an ESS, this tells the DS to void existing association information. Attempts to send messages via the DS to a disassociated STA will be unsuccessful.

- **Distribution** The AP gives a message to the distribution service of the DS. It is the job of the distribution service to deliver the message within the DS in such a way that it arrives at the appropriate DS destination for the intended recipient.

- **Integration** The Integration function is responsible for accomplishing whatever is needed to deliver a message from the DSM to the integrated LAN media (including any required media or address space translations).

- **Reassociation** The reassociation service is invoked to "move" a current association from one AP to another. This keeps the DS informed of the current mapping between AP and STA as the station moves from BSS to BSS within an ESS.

2.1.1.4 Scanning

Wireless scanning is a method to find an available wireless network access point or other nodes, in ad-hoc networks.

- **Beacons:** In an infrastructure the AP (or the nodes, in ad-hoc mode) periodically transmits special frames called *beacons* that contain a copy of it's *Timer Synchronization Function* to synchronize the other stations in a BSS. The interval between beacons is defined by a parameter of a station. The AP will define the timing for the entire BSS by transmitting beacons according to an attribute in the AP. The beacon period is included in Beacon and Probe Response frames, and STAs shall adopt that beacon period when joining the IBSS.

- **Synchronization and scanning:** A station will operate in either a Passive Scanning mode or an Active Scanning mode depending on a parameter configured therein. This scan is initialized with the reception of a primitive, which includes the type of scanning as one of its parameters. Another parameter is the SSID for which to scan. Upon reception of this primitive, the station will passively scan for any Beacon frames, or actively transmit Probe frames containing the broadcast SSID.

- **Passive Scanning:** in this case, the station will listen to each channel scanned for a duration defined in a parameter, waiting for a Beacon frame sent from a AP.
- **Active Scanning:** involves the generation of Probe frames and the subsequent processing of received Probe Response frames. A station will reply to a Probe Request with a probe response, only if the SSID in the first is the broadcast SSID or it matched the SSID of the station. Probe Response frames shall be sent as directed frames to the address of the STA that generated the probe request.

2.1.2 IEEE 802.16

IEEE 802.16 is a broadband wireless access solution for metropolitan area networks (MANs), reaching long coverage ranges and having very high throughputs, in the hundreds of Mbps range. The aim for this technology is to enable operators to reach costumers in areas where traditional wired solutions would impose high costs, and where wireless LAN reachability would not be possible. Coverage, for the IEEE 802.16 technology, can reach distances from 15 km in Non-Line of Sight environments to 50 km in Line of Sight environments.

This technology is also backed-up by a standardization effort in the form of the Worldwide Interoperability for Microwave Access (WiMAX) Forum[1]. This body's aim is to standardize equipment interoperability in order to reduce equipment prices.

The IEEE is responsible for specifying a connection-oriented Medium Access Layer (MAC) with QoS capabilities, and Physical (PHY) layers considering different frequency operations regarding Line of Sight or Non-Line of Sight operation. The IEEE also standards a set of convergence sublayers that map upper layer packets into the IEEE 802.16 system, supporting other standards such as the Internet Protocol (IPv4 and IPv6) and Asynchronous Transfer Mode (ATM). As of late, the standard now also supports mobile nodes, allowing for mobility scenarios.

2.1.2.1 802.16 Layers

The following figure shows how the 802.16 protocol stack layers look like.

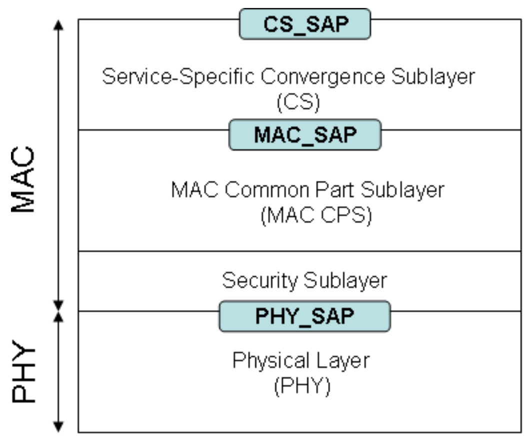

Figure 2.2: 802.16 layers

[1]http://www.wimaxforum.org

The 802.16 MAC layer is composed of three other sub-layers: the Service Specific Convergence Sublayer (CS) which is responsible for interfacing the MAC layer with higher layer entities, the Common Part Sublayer (CPS) responsible for important MAC functions such as addressing, construction and transmission of the MAC PDUs, and the Security Sublayer which is responsible for authentication and data encryption functions.

The 802.16 PHY layer supports multiple physical layer specifications customized for the frequency band of use and their associated regulations. More information can be found in the standards [11] and [12].

2.1.2.2 Basic Mode of Operation

As referred previously, the 802.16 is a connection-oriented type of technology and all packet traversal only occurs after a connection was established. Moreover, it operates under a Point-to-Multipoint mode where a special node, the Base Station (BS) is connected to a core network, and is in contact with other 802.16 nodes, the Subscriber Stations (SS). The 802.16 communication occurs between the BS and SSs, where the downlink direction is from the BS to the SS and the uplink direction is from the SS to the BS.

In the downlink direction the BS is the sole transmitter and dataframes are broadcasted to all SSs. Each SS then analyzes the Protocol Data Unit (PDU) and checks for the Connection Identifier (CID) to see if it is the recipient of said packet. In the uplink direction the uplink channel is shared amongst all SSs using a dedicated scheduling service, determining the transmission rights to each of the SSs. Four uplink scheduling services are available: Unsolicited Grant Service (UGS), real-time Polling Service (rtPS), non real-time Polling Service (nrtPS) and Best Effort (BE).

Other than the Point-to-Multipoint operational mode, there is a Mesh mode, where the SSs can communicate with each other. Also, in the 802.16e standard extension [12], the SSs can be mobile.

2.1.2.3 Connections

The establishment of a new connection between the BS and the SS assumes the existence of several pairs of management connections which are created upon SS activation. These are: the basic connection, the primary connection and the secondary connection. Each of these three connections reflects three different QoS requirements for traffic management. The basic connection is used to transfer short, time-critical MAC messages. The primary connection is used to transfer longer, more delay tolerant messages. The secondary connection is used to transfer delay tolerant, standard-based messages such as Dynamic Host Configuration Protocol (DHCP) and Simple Network Management Protocol (SNMP).

Besides these three pairs of management connections, another group of connections is also defined: the broadcast management connection, the multicast polling connection and the transport connection. The first one is configured by default and pertains to the transmission of MAC management messages to all SSs. The second connection is used by SSs to join multicast polling groups, allowing for bandwidth request via polling. The last kind of connection refers to the ones that are allocated to satisfy contracted services.

2.1.2.4 SS Network Entry and Initialization

Upon network entry the SS proceeds in the following way:

- **Scan for downlink channel and establish synchronization with the BS:** the SS scans for BS eliminating the need for manual configuration;
- **Obtain downlink and uplink parameters:** SS establishes and maintains synchronization, acquiring parameters;
- **Perform ranging:** SS tries to find an initial ranging with the BS, which answers with the CID of the basic management and primary connections;
- **Negotiate basic capabilities:** the SS and BS share knowledge of each others capabilities;
- **Perform Registration:** the SS registers with the BS, thus becoming manageable;
- **Establish IP connectivity:** the SS tries to configure an IP using DHCP via the secondary management connection;
- **Establish Time of Day:** for time-stamping management events and actions, the SS and BS must share the Time of Day;
- **Establish Provisioned connections:** the BS signals the SS to establish pre-provisioned service flows.

2.1.3 3G Systems

3G Systems are intended to provide global mobility with a wide range of services including telephony, paging, messaging, Internet and broadband data. International Telecommunication Union[2] (ITU) managed the process of defining the standard for third generation systems, referred to as International Mobile Telecommunications 2000 (IMT-2000). In Europe the European Telecommunications Standards Institute (ETSI) was responsible for the Universal Mobile Telecommunications System (UMTS) standardisation process. In 1998 the Third Generation Partnership Project (3GPP) was formed to continue the technical specification work. 3GPP has five main UMTS standardisation areas: Radio Access Network, Core Network, Terminals, Services and System Aspects and GSM-EDGE Radio Access Network (GERAN).

The 3GPP Radio Access group is responsible for tasks such as specification of radio layers and interfaces, UMTS Terrestrial Radio Access Network (UTRAN) operation and maintenance requirements, radio performance specification, conformance tests and specification for radio performance aspects from the system point of view.

The 3GPP Core Network group is responsible for defining the mobility management and call connection control signalling between the user and the core network, the signalling between the core network nodes, the interworking functions between the core and external networks and packet related issues.

The 3GPP Terminal group is responsible for defining service capability protocols, interworking of end-to-end services, servicing and interfacing with the mobile terminal and conformance tests specifications of terminals.

The 3GPP Services and System Aspects group is responsible for the definition of services and features requirements, developing service capabilities and architectures for applications (cellular, fixed and cordless), charging and accounting, network management, security aspects and maintenance of the overall architecture.

[2] http://www.itu.int

The standardization effort has generated numerous documents pertaining to technical and applicational specifications. More information regarding specific subjects can be found on 3GPP documents such as: [13] for network architecture, [14] for handover procedures, [15] for group services and system aspects and [16] for WLAN interworking.

2.1.3.1 UMTS Services

UMTS offers teleservices (like speech or SMS) and bearer services, which provide the capability for information transfer between access points. It is possible to negotiate and renegotiate the characteristics of a bearer service at session or connection establishment and during an ongoing session or connection. Both connection oriented and connectionless services are offered for Point-to-Point and Point-to-Multipoint communication.

Bearer services have different QoS parameters for maximum transfer delay, delay variation and bit error rate. Offered data rate targets are:

- 144 kbits/s satellite and rural outdoor;
- 384 kbits/s urban outdoor;
- 2048 kbits/s indoor and low range outdoor.

UMTS network services have different QoS classes for four types of traffic:

- Conversational class (voice, video telephony, video gaming);
- Streaming class (multimedia, video on demand, webcast);
- Interactive class (web browsing, network gaming, database access);
- Background class (email, SMS, downloading).

UMTS will also have a Virtual Home Environment (VHE). It is a concept for personal service environment portability across network boundaries and between terminals. Personal service environment means that users are consistently presented with the same personalised features, User Interface customisation and services in whatever network or terminal, wherever the user may be located. UMTS also has improved network security and location based services.

2.1.3.2 UMTS Architecture

A UMTS network consist of three interacting domains; Core Network (CN), UTRAN and User Equipment (UE). The main function of the core network is to provide switching, routing and transit for user traffic. The core network also contains the databases and network management functions.

The basic Core Network architecture for UMTS is an evolution of the Global System for Mobile Communications (GSM) network with General Packet Radio Service (GPRS).

2.1.3.3 Core Network

The Core Network is currently divided in circuit switched and packet switched domains. Some of the circuit switched elements are Mobile services Switching Centre (MSC), Visitor location register (VLR) and Gateway MSC. Packet switched elements are Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). Some network elements, like EIR (Equipment Identity Register), HLR (Home Location Register), VLR (Visitor Location Registry) and AuC (Authentication Center) are shared by both domains.

The Asynchronous Transfer Mode (ATM) is defined for UMTS core transmission. ATM Adaptation Layer type 2 (AAL2) handles circuit switched connection and packet connection protocol AAL5 is designed for data delivery.

The architecture of the Core Network may change when new services and features are introduced. Number Portability DataBase (NPDB) will be used to enable users to change the network while keeping their old phone number. Gateway Location Register (GLR) may be used to optimise the subscriber handling between network boundaries. MSC, VLR and SGSN can merge to become a UMTS MSC. It is necessary for a network to know the approximate location in order to be able to page user equipment. System areas can be (from largest to smallest) UMTS systems (including satellite), Public Land Mobile Network (PLMN), MSC/VLR or SGSN, Location Area, Routing Area (PS domain), UTRAN Registration Area (PS domain), Cell or, lastly, Sub cell.

2.1.3.4 Radio Access

Wide Band Code Division Multiple Access (WBCDMA) technology was selected for the UTRAN air interface. UMTS WCDMA is a Direct Sequence CDMA system where user data is multiplied with quasi-random bits derived from WCDMA Spreading codes. In UMTS, in addition to channelization, Codes are used for synchronisation and scrambling. WCDMA has two basic modes of operation: Frequency Division Duplex (FDD) and Time Division Duplex (TDD).

A Base Station is referred to as Node-B and the control equipment for Node-B's is called Radio Network Controller (RNC). The functions of a Node-B are transmission and reception of air interfaces, modulation/demodulation, CDMA physical channel coding, micro diversity, error handling and closed loop power control.

The functions of a RNC are controlling the radio resources, admission control, allocating channels, setting power controls, handover control, macro diversity, ciphering, segmentation/reassembly, broadcast signalling and open loop power control.

2.1.3.5 User Equipment

The UMTS standard does not restrict the functionality of the User Equipment in any way. Terminals work as an air interface counterpart for the Node-B and have many different types of identities. Most of these UMTS identity types are taken directly from GSM specifications.

- International Mobile Subscriber Identity (IMSI);
- Temporary Mobile Subscriber Identity (TMSI);
- Packet Temporary Mobile Subscriber Identity (P-TMSI);
- Temporary Logical Link Identity (TLLI);

- Mobile station ISDN (MSISDN);
- International Mobile Station Equipment Identity (IMEI);
- International Mobile Station Equipment Identity and Software Number (IMEISV).

UMTS mobile stations can operate in one of three modes of operation:

- **PS mode of operation:** The MS is attached to the PS domain only and may only operate services of the PS domain. However, this does not prevent CS-like services to be offered over the PS domain (like VoIP);
- **CS mode of operation:** The MS is attached to the CS domain only and may only operate services of the CS domain.
- **Packet Switched (PS) / Circuit Switched (CS) mode of operation:** The MS is attached to both the PS domain and CS domain, and the MS is capable of simultaneously operating PS services and CS services;

The UMTS Integrated Circuit Card (UICC) has the same physical characteristics as GSM Subscribed Identity Module (SIM) card. It has several functions such as supporting User Service Identity Module (USIM) applications, user profiles of the USIM, update USIM specific information over the air, security, user authentication, payment methods and secure downloading of new applications.

2.2 Multimode Terminals

Over the past few years, the range of wireless network services and technologies has increased rapidly. In addition to this, user expectations regarding what services they would like to access whilst on the move have become more demanding. For example, users are beginning to demand bandwidth intensive applications, such as real-time movie players, but want to pay as little for this service as possible.

In response to this, the next generation of mobile terminals will include multiple interface technologies, allowing the reception of data over multiple system bearers, with different characteristics. For example, GPRS provides country wide coverage, but with restricted bandwidth and comparatively high cost per bit, whereas WLAN coverage provides low cost, high bit rate services, but with restricted coverage. By integrating both these technologies into a single device, it is possible to support the establishment of application sessions across whichever interface best meets the application requirements and user preferences, and to dynamically switch this application between interfaces as the user moves in and out of the vicinity of a particular technology.

A number of requirements for multimode terminal operation include:

- Operating with minimal inputs from the user. From a user experience perspective, it is preferable to carry out these decisions in an automated way without having to query the user every time a new interface becomes available, or an old interface disappears.
- Selecting interfaces based on user preferences, application requirements and information about the network
- Balancing traffic over available interfaces in a way that is transparent to the user, i.e. as seamlessly as is possible.

2.2.1 Deployment Issues

To support intelligent multimode terminal operation, two issues must be considered:

- **Network Detection:** The availability of connectivity must be detected by the terminal. For some interface technologies the resources used to maintain the interface powered on, even when not active, are relatively small and the interface is able to stay permanently on and scan for coverage. In the case of IEEE 802.11 WLANs, this approach is unfeasible due to the amount of power needed to keep the interface actively scanning for access points. So, the question arises as how to activate the WLAN interface when in the vicinity of a hotspot. Currently, this activation of the interface is left up to the user, who manually enables their WLAN card when in a hotspot. However, since one of the noted requirements was to minimise user interaction, a desirable goal might be to automate this process. One solution that would achieve this result would be to advertise the presence of nearby services (or coverage of other technologies) via the currently active interface. For example, an indication of the presence of a hotspot could be sent via GSM, which would then activate the WLAN interface automatically.

- **Network Capability and Characteristic Discovery:** One key issue that has become apparent recently with the proliferation of different link layer technologies is how service providers can offer a consistent service across heterogeneous networks (and indeed across access networks that they do not own themselves). From an operator perspective, they want to present a single, easily understood brand to their costumers, but they have to have many different variants for different layer 2 technologies. From a terminal perspective, the decisions made by Interface Selection and Session Transfer Management are only as good as the information made available to them. Whilst configuring the terminal with local preferences and application requirements is fairly simple, finding out information about the network to support interface selection decisions needs some additional solutions.

2.2.2 Multihoming

A computer system is denoted as multihomed when it has multiple connections to physically separated networks, either wired of wireless. Also, computer systems with only one physical interface, but with several configured global IP addresses, can be seen as multihomed.

2.2.2.1 Types of Multihoming

Several types of multihoming exist:

- **Host Multihoming:** when a host has multiple physical network interfaces, with the same or different access technologies, getting access through them at the same time is known as host multihoming. At least one IP address per network interface has to be configured.

- **Site Multihoming:** when a host is connected to a site supplied by two or more access routers, is known as site multihoming. The terminal can send and receive packets through each one, provided it has an IP configured for each of the access routers. So, the terminal can choose which access router (which, in turn can belong to difference ISPs) to use to communicate, or even balance the communication.

2.2.2.2 Usage Scenarios

Multihoming is commonly used as a mean to apply load balancing, increase available bandwidth, provide connectivity redundancy and disaster recovery. Multihoming allows for improved coverage and the ability to switch between different types of access, when more than one is available. Through the use of MobileIP, it is possible to maintain a session when switching between accesses. Examples of scenarios in which multihoming can be used are:

- **Increase Bandwidth:** bandwidth can be increased through the use of load balancing between multiple interfaces, wired of wireless. Multiple flows can be distributed across the interfaces, or a single flow can be split across the interfaces.

- **Increase Availability:** a change of interface can be done in case the main one suffers a failure, securing connectivity.

- **Improve QoS:** a change of interface can be done in case the main one suffers from QoS deterioration, providing better connection conditions.

- **Transfer Cost Optimization:** selection of the best interface considering flow type.

- **Reduce packet loss on slow Handoffs:** usage of multiple interfaces in parallel to provide sustained connection during a slow handoff to another interface.

- **Routing/Gateway functionality with multiple interfaces:** a user with multiple interfaces can share, for example, one of his connections with nodes found through other connections, through IP forwarding.

The following items depict some considerations exploring network approaches to achieve multihoming.

- **Mobility Support** In a multihomed environment, a mobile terminal owns more than one IP address. MIPv6 allows the use of such addresses as CoA, but the protocol specification only allows a single CoA to be registered with the HoA. To achieve a multihoming environment, the mobile terminal should consider all configured IP addresses as CoAs and advertise them on the Binding Update messages, thus binding the different CoAs to a single HoA. Each CoA could be bound with many HoAs and vice versa. Those bindings are chosen using some kind of policying like bandwidth requirement, network load, delay, etc. The transport layer uses the HoAs to establish a new session: if the terminal is in the Home Network, the chosen HoA is used by the network layer for the communication. When the mobile terminal is in a foreign network the transport layer still uses the HoA for the communication, but MIPv6 rewrites the data packets using one of the CoA associated to the HoA.

- **Routing** One approach to achieve multihoming is the routing protocol, extending it to IPv6. The aim of this approach was to avoid modification to the hosts, which was achieved through the assignment of a global prefix to the multihomed site. This prefix is, in turn, advertised within the network by routing protocols of the networks to which the multihomed site is connected to. From the host point of view, a device connected to a multihomed site owns only one locator which also identifies it within the Internet. Nevertheless this approach gives some concerns regarding scalability and the time required by the routing protocol to update the routing table after a topological change.

- **SHIM6** SHIM6 is a protocol based on modifications at the IP level. This protocol is defined to achieve site multihoming, mainly preserving established communications through failures. At the start of a session, the mobile terminal and the correspondent node exchange their list of available IP addresses. The session is started with the choice of one of the IP addresses. If this IP address becomes no more usable, the SHIM6 layer chooses another of the available IP addresses and tests its usability. When a usable IP address is found, it changes the packets' headers to comply with the new conditions. Nevertheless, the SHIM6 is still under definition and it only covers the challenge of reliability. Also it is not designed to solve problems such as host mobility.

- **3GPP** As IPv6 allows to have multiple addresses in each interface, the 3GPP terminal with a single interface can be multihomed using multiple IPv6 addresses with a single radio connection. Furthermore some 3GPP terminals are available with more than one interface (3GPP and WLAN) can be multihomed using one or several IP addresses in each interface. Similarly, a GGSN can have multiple IPv6 addresses per interface and several upstream links.

2.3 IP Mobility

Discussions on heterogeneous networks agree on the need of a common protocol for communication, the IP protocol. IP Mobility stands as the convergence point at which different technologies are able to be used to allow a terminal to maintain its ongoing session. This section presents concepts and new directions related to this subject.

2.3.1 Mobility Related Terminology

Handover terminology is related to different perspectives and approaches to the several aspects of mobility. Distinctions can be made according to the scope, range overlap, performance, diversity, state, mobility types and control modes of handover techniques.

Handover, also called handoff, is a process by which an active MN changes its point of attachment to the network, or when such a change is attempted.

2.3.1.1 Scope

The scope of a handover can be classified as:

- **Layer 2 Handover** This handover is characterised when the MN changes APs connected to the same AR's interface. This kind of handover is transparent to the routing at the IP layer.

- **Intra-AR Handover** A handover which changes the AR's network interface to the mobile. That is, the Serving AR remains the same but internal routing changes to the AR take place.

- **Intra-AN Handover** A handover where the MN changes ARs inside the same AN. This handover may not be visible outside of the AN. This happens in case of a change of AN Gateway, due to a change in the MN's data flows.

- **Inter-AN Handover** A handover where the MN moves to a new AN. There is a requirement for macro mobility support concerning this handover, as well a possible new IP address to the MN.

- **Intra-technology handover** A handover between equipment of the same technology.

- **Inter-technology handover** A handover between equipment of different technologies.

- **Horizontal handover** A handover which involves MNs moving between access points of the same type.

- **Vertical handover** A handover where the MNs are moving between access points of different types.

Inter-technology handovers are of special importance due to the subject of this thesis. A mobile node may be equipped with multiple interfaces, where each interface can support a different access technology (802.11, CDMA). A mobile node may like to communicate with one interface at any time in order to conserve power. During the handover the mobile node may move out of the footprint of one access technology (e.g., 802.11) and move into the footprint of a different access technology (e.g.,CDMA). This will warrant switching of the communicating interface on the mobile as well. This type of Inter-technology handover is often called Vertical Handover since the mobile node moves between two different technology cells, with different sizes. A vertical handover can be termed as upward vertical handover or downward vertical handover based on the direction of movement such as smaller cell to larger cell or vice versa. A mobile node moving from 802.11 network to cellular network can be viewed as upward vertical handover. An inter-technology handover may affect the quality-of-service of the multimedia communication, since each access network offers different bandwidth.

One has to note that horizontal and vertical handovers might not be noticed at IP level. For instance, an AR could control WLAN and Bluetooth access points, and the mobile node could do horizontal and vertical handovers under the same AR without changing its IP address or even the network interface.

One of the major issues involving heterogeneous handovers is the handover delay. As a mobile node goes through a heterogeneous handover process, it is subject to handover delays because of the rebinding of properties at several layers of the protocol stack. This delay mostly pertains to factors such as access characteristics (e.g., bandwidth, channel characteristics), access mechanisms (e.g. CDMA, CSMA/CA, TDMA), configuration of layer 3 parameters, re-authentication, re-authorization, rebinding of security association at all layers, etc.. [17] proposes some means to optimize handover delay such as performing the authorization and authentication phases prior to the handover.

2.3.1.2 Control

Handover control takes place at instances such as origin of the handover, primary control and assistance. A handover must either be **Mobile-initiated Handover** (MIHO), when the initial decision for handover starts in the MN, or **Networked-initiated Handover** (NIHO), when it is the network who makes the initial decision to initiate the handover. Also, the handover can be **Mobile-controlled** when the MN is the primary controller of the handover, or **Network-controlled**, when that role falls to the network. The measurements made that indicate when and where to target the handover are also aspects of distinction of handovers. So, depending on the occurrence of such measurements, a handover can be assisted or not. In case of a assisted handover, it can either be a **Mobile-assisted** handover, when information and measurement from the MN are used by the network, or a **Network-assisted** Handover, when such details are collected by the network and sent to the terminal. A handover can also be considered a **Push** handover, when the previous network attachment point is involved in the handover initiation, of a **Pull** handover, when the new network attachment point is involved in the handover initiation. Lastly, a **Planned** handover is a proactive handover where signalling was done in advance, and an **Unplanned** handover is a reactive handover where no signalling was done in advance.

In [18] a problem statement on network-initiated handovers is presented where new potential deployment scenarios, and associated functionalities, are considered regarding support for decision making and execution of handovers. One can regard the addition of network control to local and global mobility domains, as a way for mobile operators to gain additional control on user mobility while taking into account the allocation of resources and more reliable end to end services guarantee. Application areas such as Mobility (for handling user mobility from the network side), Resource Optimization (for network optimization) and Inter-Domain Handovers (when crossing administrative domain borders) can be considered.

Also in [19] challenges and possibilities on NIHO scenarios are analyzed. In complex diverse heterogeneous network scenarios it is often difficult for the mobile node to have enough information, or even the possibility, to execute intelligent handover decisions. In these multi-access scenarios, preferred network selection is not necessarily based on access availability but rather on policies and commercial roaming agreements at access network, access provider and service provider levels. Also, this information can be dynamic, changing periodically making it impossible to store in the mobile node, and some of the information is only available to the network. Radio resource management information, which resorts to mechanisms and algorithms able to gather measurements and force terminal handover between cells, is an example of information only available to networks. This network information ultimately leads to better service provision to users, additionally providing the network operator with increased resource usage: flow balancing, optimized resource re-allocation, mechanisms for user traffic performance optimization and mechanisms for user profiling. This document also references two kinds of network initiated handovers: "**Opportunistic Handovers**", performed for resource optimization; and "**Essential Handovers**", performed for link layer condition changes such as degradation of the received signal level.

The network needs to be aware of the mobile terminal view of the network, in order to execute an informed and optimized network initiated handover. On NIHO scenarios, one of the major concerns is the freshness of information conveyed to network from the mobile node. Terminal mobility and channel interference can impact the accuracy of the conveyed information. Thus, considering these issues, the serving radio entity, upon measurement reports, can request a particular mobile device to move to another point of attachment.

This serves as the background approach for studying how MIH mechanisms can support and optimize network initiated handovers in heterogeneous networks, described in chapter 5.

2.3.1.3 Connectivity

Regarding simultaneous connectivity to Access Routers, a handover can be **Make-before-break** of **Break-before-Make**. In the first case, the MN makes the new connection before the old one is broken. So, the MN can communicate simultaneously with the old and new AR during the handover (i.e., multimode terminals). The second case occurs when the MN breaks the old connection before the new one is made.

2.3.1.4 Performance

Performance and functional aspects are important for handover classification. For example, handover latency is referred as the difference between the time a MN is last able to send and/or receive an IP packet by way of the previous network link, and the time the MN is able to send and/or receive an IP packet through the new network link. So, a **Fast handover** is a handover that aims primarily to minimize handover latency. This type of handover, however, incurs no explicit interest in packet loss. Under this concept, **Smooth handovers** are the handovers that aim primarily at minimizing packet loss. A **Seamless handover** is defined as a handover in which there is no change in service capability, security, or quality. In practice, some degradation in service is to be expected, but it might not be perceptible or affect high layer applications. As a consequence, what would be a seamless handover for one less demanding application might not be seamless for another more demanding application.

2.3.1.5 Types of mobility

- **Macro mobility or Global mobility** Related to mobility over a large area. This includes mobility support and associated address registration procedures that are needed when a MN moves between IP domains. **Inter-AN** handovers typically involve macro-mobility protocols. Mobile-IP can be seen as a mean to provide macro mobility;

- **Micro mobility or Local mobility** Occurs over a small area. Usually this means mobility within an IP domain with an emphasis on support for active mode using handover, although it may include idle mode procedures also. Micro-mobility protocols exploit the locality of movement by confining movement related changes and signalling to the access network.

2.3.2 Common Mobility Approaches and Enhancements

IP mobility (namely Mobile IP [20]) provides Internet connectivity to mobile nodes roaming from one access router to another, regardless of the access technology supported in the router. Mobile IP is based on the existence of a Home Agent, the creation of a Care Of Address when roaming, and the establishment of tunnels and/or specific route update mechanisms that reroute the traffic from the home to the visited network. To reduce handover latency and increase scalability, extensions have been developed (e.g.[21], [22]). Although some of these approaches [21] may potentially support network initiation of the handover procedure, none of them takes into consideration the existence of a backend combining mobility, resource management and roaming scenarios. Previous work [23] shows the advantages of having a mobile initiated handover considering information provided by the network, in wireless LAN environments where HMIPv6 is implemented. Both implementation and simulation results are provided, however the system considers only wireless LAN, and handovers are triggered only by terminal mobility. In [24], the IP2 Mobility Management Protocol was developed to satisfy mobile operator's requirements for next-generation mobile networks. IP2MM exploits the network controlled approach instead of a terminal controlled mobility management. The main conclusion is that with this type of protocols a better performance of the network can be achieved but the paper mostly addresses the performance in mobility control and packet throughput.

However, it should be noted that none of the mentioned approaches consider a scenario in which mobile initiated handovers (MIHO) are combined with network initiated handovers (NIHO). 4G networks will require this combination, allowing personalization in the user's terminal (network preferences) and resource usage optimization by the network. Also, the dynamism, cell coverage, and multi-technology environment is different from the traditional scenario of cellular networks, thus the results of network initiated handover in those networks are not directly applicable to 4G networks. In [25] initial simulation results and implementation experiences prove the feasibility of the 4G approach covering a wide range of access technologies.

2.3.3 Extending Mobility-based Architectures: Localized Mobility

Mobile IP [20] allows for mobility support at the IP level and is becoming intrinsically supported in IPv6 (or with novel proposals such as HIP [26]). MIP has nevertheless well-known deficiencies both in terms of performance and functionalities. Thus most of the research being done recently has been focused in these aspects, in particular along the lines of localized optimization of mobility behavior, separating the local mobility from the global mobility (MIP-based).

The localized mobility proposals aimed initially to reduce signaling outside the local domain, and improve efficiency by managing the "local" mobility closer to the mobile node MN (reducing the time needed for the signaling required by the mobility). Recently, operational exploitation considerations are gaining increasing relevance.

2.3.3.1 Host Based Localized Mobility Management

Initial localized mobility techniques were host-based, i.e. hosts had to handle signaling, and to be aware of local and global signaling protocols. Most relevant previous protocols were Hierarchical MIP and Cellular IP.

Hierarchical Mobile IPv6 [22] is a protocol that extends Mobile IPv6 by introducing a dedicated device, named Mobility Anchor Point (MAP), whose task is to handle the movement of a MN within a defined set of Access Routers (AR). Thus the MAP handles local mobility while Home Agent (HA) handles movements among different local domains. The introduction of this hierarchy (i.e. splitting of mobility management between HA and MAPs) aims at optimizing overhead during handovers among ARs of the same local domain: signaling control is reduced since it is exchanged only between Mobile Node (MN) and MAP and therefore handover latency is reduced. HMIPv6 is a host based solution which requires MN to control both local domain and global domain signaling. Another drawback of HMIPv6 is that it introduces an additional tunnel over the air.

Cellular IP (CIP) [27] is slightly different. MN registers as its CoA the IP address of a node (called Gateway) in the LMD. So all the traffic addressed to the MN will reach the Gateway. To send the traffic from the Gateway to the MN, host routes associated with the HoA of the MN are used inside the LMD. To create these host routes, the (non standard) routers of the LMD use the uplink traffic of the MN (or special purpose signaling packets) to update or refresh the route in the reverse direction. Each router will learn from the uplink traffic which is (for the downlink traffic) the next hop to which to send the packets destined to the MN. All uplink traffic is forwarded hop-by-hop to the Gateway regardless of its destination. Cellular IP is a host based solution, as the host has to register explicitly in the Gateway when it arrives to the LMD. This allows the MN to discover the CoA that it must register, and it allows the Gateway to learn that it must start forwarding packets to/from the MN. Also route updates may require the MN to send special purpose packets.

2.3.3.2 Network Based Localized Mobility Management

Aspects of network control and operation have led to the renewed development of localized mobility solutions, particularly at standardization level in the IETF. Unlike host-based mobility where mobile terminals signal a location change to the network to achieve reachability, network based approaches relocate relevant functionality for mobility management from the mobile terminal to the network.

The *NetLMM* [28] approach was designed in the IETF NetLMM Working Group. [28] and [29] define the requirements and rational for NetLMM. The network learns through standard terminal operation (such as router and neighbour discovery or by means of link-layer support), about a terminal's movement and coordinates routing state update without any mobility specific support from the terminal. This approach allows hierarchical mobility management: mobile terminals signal location update to a global mobility anchor only when they change the LMD and, in the LMD, mobility is supplied to terminals without any support for mobility management. NetLMM complements host-based global mobility management by means of introducing local edge domains.

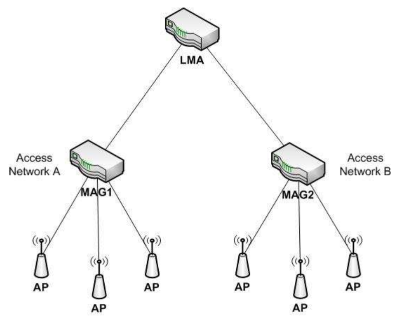

Figure 2.3: NetLMM domain

Figure 2.3 depicts the entities involved in a NetLMM architecture. The Local Mobility Anchor (LMA) is a router defining the edge between the NetLMM domain and the core network. If a global mobility scheme is used, it is the boundary between GMD and LMD. The Mobility Access Gateway (MAG) is the Access Router for the MN. Note that the host routes are only configured in the LMA and the MAG; all intermediate nodes are totally NetLMM unaware, which considerably reduces the signalling in the LMD and avoids the extensive use of resources (routing tables) in the intermediate nodes (which is not the case with Cellular IP, for example).

The NetLMM area of operation is located between the LMA and the MAG. The forwarding method used between the MAG and the LMA (IPv6 in IPv6 tunnelling) can be extended to General Routing Encapsulation [30] or Multi Protocol Label Switching [31].

The NetLMM protocol builds an overlay network on top of a physical network where the terminals are capable to roam across MAGs without changing IP address configuration. The basic protocol defines only reactive handover and does not consider the support for multiple technologies within the same LMD. Both functions, enabling proactive handovers across the different technologies are required.

More recently, attentions have been turned to another localized mobility protocol, Proxy Mobile IPv6 (PMIPv6). According to [32] PMIPV6 is a network-based mobility management protocol aimed at local mobility support, while reusing when possible Mobile IPv6 [20] entities and concepts. In this protocol the mobile nodes are differentiated by a network address identifier (NAI), with an associated set of information stored in the network, such as a profile containing the home prefix. This information is typically kept in a policy store (e.g. AAA), accessible by all the PMIP entities in the local mobility domain.

PMIPv6 assumes that upon L2 network attachment, the node is authenticated. This attachment provides the necessary information (e.g. the nodes NAI) to ensure that the network is able to retrieve the Home Network prefix. The prefix will then be used in Router Advertisements to the node, informing that it is on the Home Domain. In this scenario the MN configures its Home Address on the network interface, even when roaming across foreign networks, transforming the visited LMD into a single link, from the node's point of view. The Proxy Mobile Agent (PMA), located in the access router, performs signaling on behalf of the node and is also the entity that retrieves the MN information and sends the customized Router Advertisements, emulating the home network behavior. The PMA mobility signaling consists on Binding Updates to the MN's Home Agent, informing the HA that the current Care-of Address of the registered MN is the PMA's address. These procedures also lead to the establishment of tunnels between HA and PMA.

The PMIPv6 concept has also been applied to the NetLMM concept presented earlier. The Protocol for Network based Localized Mobility Management (PMIP-LMM), as described in [33] , includes several entities: the MAG, the Proxy MIP Client (PMIP-Client) and the LMA. The behavior, shown in figure 2.4, differs from the previously presented solution in both operation and addressing.

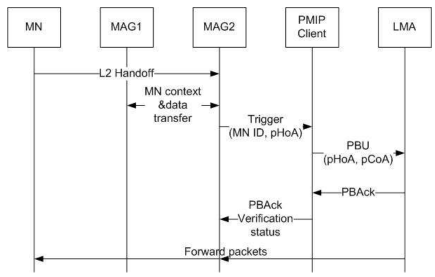

Figure 2.4: Message flow for PMIPv6 concept applied to NetLMM

While on visited networks the MN acquires a proxy Home Address (pHoA), which is the MN's address on the LMD. The acquired pHoA is in fact the CoA registered at the Home Agent, in Binding Update procedures. The address representing the MN's location inside the LMD is the Proxy Care-of Address (pCoA), which corresponds to the MAG's address, currently serving the MN. Upon L2 attachment of the MN to the LMD, the MAG obtains the MN's Link Layer ID. The MAG than sends a Router Advertisement to the node, allowing it to configure the pHoA, which can be used on to update the Home Agent. With the pHoA and the Link Layer ID, the MAG registers the MN on the PMIP Client, storing the current pHoA and the pCoA, which are later used in Proxy Binding Updates (PBU) to the LMA, creating the necessary tunnels for packet delivery inside the LMD.

2.4 Towards a 4G Heterogeneous Mobile Network Environment

4G networks do not represent specific technologies but rather a vision and a market concept enabled by different technologies. From a high-level perspective, 4G networks [34] rely on three main principles:

- 4G networks are implemented in such a way that the functions performed by the network are separated into functional planes. These include access, transport, control & intelligence, and service. Each layer is independent of the others in terms of modifications and upgradability. This architecture provides a flexible and scalable network, reducing time to market for the implementation of new services.

- The functional planes are separated by open interfaces that allow interconnection to other operator's networks but also the integration of third-parties's services and applications. Provided that commercial agreements are established between the different parties, this principle can widen the operator's coverage and service scope, providing end-users with access to a greater number of services.

- 4G networks are multi-service networks, as opposed to legacy networks that are only used for specific services. This kind of network allows operators to implement converged and new services. From the user's perspective, the convergence of services will enable the emergence of the seamless service concepts, where users can access their "home" services from any type of access network.

2.4.1 General Architecture for 4G Networks

As stated before, the architecture is composed of functional planes that perform tasks at different levels:

- The **access plane** provides the infrastructure, i.e. the access network, between the end-user and the transport network. The access plane may be wireless or wireline, and it can be based on different transmission media e.g. wireless, cooper wires, cable and fibre optic. Technologies in the access plane can be circuit-switched or packet-switched. The access network is connected to network nodes at the edge of the backbone network.

- The **transport plane** provides transport between network nodes to which the access networks are connected. The transport plane consists of one or several backbone networks based on packet or cell switched network nodes. Links are mainly based on optical fibre links but can also be satellite or terrestrial links. The transport plane is capable of handling different kinds of traffic, e.g. voice conversation, streaming video, interactive data, and batch data. Gateways at the edge of the transport network convert traffic to and from legacy networks, e.g. telephony, Internet and real-time data applications.

- The **control plane** includes network control elements. As such, the control plane controls all other planes. The control plane can, for instance, be responsible for the control of communication sessions, e.g. establishing or disconnecting voice calls or multimedia sessions, intelligent service provisioning or resources provisioning depending on the service requested. A core principle of the 4G architecture is to separate the control logic from the underlying switching hardware.

- The **service plane** offers elementary service functions that can be used by service providers to build more complex or comprehensive services. The elementary functions are related to transport, traffic, content and service AAA & billing. The service plane also provides interfaces towards service providers who want to use these elementary service functions to access the underlying infrastructure. Such access will depend on commercial agreements between service providers/third parties and network operators. The interfaces may be implemented in different ways, e.g. in the form of APIs for service specific software to be run on servers within the network, or in the form of open standardised interfaces between the network and application servers. Such interfaces will enable the unbundling of services and underlying technologies.

The demands on telecommunication networks today and in the future are reflected clearly in the societal changes around the world. Customers of all services have historically wanted more for less, more flexibility, and sometimes just "more"! This can certainly be seen in the rapid growth of mobile network usage. The mentality of anytime, anywhere is becoming prevalent around the world, the primary differences between regions being largely a matter of degree. Also, customers are seeing innovation across all areas of consumer goods and services with "just in time" services tailored to their specific needs and timeframes. These too will be key demands placed upon telecommunications networks as they evolve.

Key characteristics of 4G networks are:

- **Geographic transparency:** boundaries are disappearing and economic benefits independent of service "density" must be realized;

- **Transport efficiencies:** transport costs (price/bit) are continuously declining, 4G networks must share these efficiencies for both bearer and signaling traffic;

- **Internet technology economics:** leverage services and service delivery through the Internet, as well as the "silicon economics" of Internet hardware (servers, etc.) as memory and processor price/performance improve;

- **"Old World" to "New World" interoperability:** existing PSTN infrastructure, and its associated investment must be fully utilized.

The development of access technologies is an important driver for the development of 4G networks and will impact the development of broadband services and content. 4G networks will rely on a number of heterogeneous networks of different size and different coverage linking together servers, workstations and terminal devices, as the Internet does today. In the long term, IP will be the common interworking technology for 4G networks.

If 4G networks are to provide convergent services and to become multi-service networks, a number of interworking issues between networks has to be overcome, such as routing and QoS.

In the service convergence environment of 4G networks, one could also use the same types of tools and the same technologies to implement any type of services, e.g. telephony services or Web Services. A number of standards have already proven to be successful in this area and will play an important part in the 4G migration: Java, standardised APIs such as JAIN, Parlay and OSA, XML, etc.

A direct consequence of the emergence of such standards and technologies is that an 4G environment offers opportunities for third party service providers. In traditional telco networks such as PSTN and GSM networks, the services accessible by the end-users are limited by what has been implemented by the operators on their networks, most often by using proprietary development tools. Users can "control" these services (like for instance by enabling or disabling call barring) but only within the scope of what has actually been allowed by the operator. In an 4G environment based on IP and the tools and technologies described above, third parties can have a greater importance and the end-users will also have more choice with regards to services and environment personalisation - provided that there are not exclusionary practices and that 4G networks are open to the third-party service providers.

2.4.2 Network-Initiated Handovers Centralized Framework

Traditional Mobile Initiated Handovers are triggered by mobile devices upon collection of events such as radio signal level degradation, application requirements or the like. MIHO mechanisms can be further improved by retrieving from the access network information [35] about available bandwidth, network load in a specific access point/access router, etc.. However, disclosing such information to mobile devices is subject to access network policies and might not always be possible to provide such data. It would be desirable to gather information in the access network about load conditions (in a network-to-network relation) as well as from mobile devices (in a mobile to network relation) leading to the composition of an accurate and dynamic map that handover decision engines (located in the access network) could benefit of. Hence, NIHO via a centralized approach aims at improving network operations where required. The concept of MIHO and NIHO applies to both intra-technology case or inter-technology case, the former being potentially layer two specific, the latter leading to an IP based approach. A simulation study comparing MIHO and NIHO performance can be found in [19], where it is shown that NIHO provides improved resource allocation when heterogeneous wireless/wired access technologies are deployed. It also quantifies the benefits of the NIHO approach and identifies conditions that affect the relevance of NIHO support.

A set of functionalities and associated protocol operations are required to support, in a common framework, MIHO and NIHO. An architecture should specify modules implemented in the terminal side, in the access part of the network and in the home domain. The mobile device should offer a cross layer design for wireless/wired technology management, compounded by an intelligent module for interface selection based on several parameters spanning from layer two related information up to user preferences and context aware applications. Protocol communications between the mobile device and the access router (first layer-3 hop in the access network) should also allow exchange of information for neighbor discovery, handover preparation and handover execution. Handover target selection is done through the combination of mobility and resource management mechanisms such as admission control. That is, handovers could potentially be denied to mobile devices. To improve performance and avoid as much as possible these situations the architecture should offer the possibility to initiate handovers from the network upon, for instance, load detection or conditions changed due to the availability of new access technologies.

2.4.3 The DAIDALOS case

Some of the previously presented concepts have already been demonstrated in the framework of several European founded projects. One of those is the DAIDALOS[3] project, standing for Designing Advanced Network Interfaces for the Delivery and Administration of Location Independent, Optimised Personal Services. This project aims to successfully demonstrate the feasibility of a platform where mobility, security and quality of service are tightly integrated, considering advanced mobility procedures looking at technology specific requirements as well at new mobile operator requirements, under an heterogeneous environment.

Daidalos exploits the so-called MARQS concept. This concerns the integration of **M**obility Management, **A**AA, **R**esource Management, **Q**oS and **S**ecurity aspects. The interaction of mobility procedures with authorisation and authentication, quality of service provisioning and resource control is of fundamental importance in mobile systems. In this view, mobility management assumes a novel and fundamental role. The architecture recognizes the current trend in networks deployment to a heterogeneous landscape of access providers [28], [15]. In such an environment it is important to give to the access providers (e.g. ISP or Network Access Point (NAP)) the flexibility for managing users mobility inside their own domain without requiring an interaction with the mobile operator domain. Thus, it is envisioned the separation of mobility management [36] into different levels: a global level associated with the Mobile Operator network and a local level associated to network access providers. The key aspect of this partitioning is the association of mobility management at administrative domain. This brings considerable flexibility to operators allowing each one to choose their preferred methods without being dependent of a particular scheme. The management of the mobility in these two levels is kept completely independent, this being a key characteristic of the architecture, absent in the traditional hierarchical mobility management approaches.

The architecture addresses as well the case where the local mobility management domain is implemented either at layer three or at layer two. The advantage of making the local mobility management

[3]http://www.ist-daidalos.org

independent of the global allows the access provider to implement legacy technologies such as WiMax architectures [12], 3GPP architectures [15] as well as layer two technologies such as IEEE 802.11r. To allow easy integration with the terminal side it is envisioned the specification of a single interface abstracting the communication with the local mobility management scheme. On the one hand the terminal always generates the same set of triggers/events and responds to the same set of commands independently of local mobility management scheme. On the other hand the network manages user mobility according to its own scheme transparent to the end user.

Mobility management is tightly coupled with Resource Management and QoS aspects focusing therefore on the MRQ of the MARQS concept. To efficiently support Resource management and QoS the architecture proposes the development of a single framework for signalling. Starting from the upcoming standard IEEE 802.21 [1] a new set of functionalities and extensions is being investigated to provide heterogeneous support for network/mobile initiated handover (NIHO/MIHO) as well as optimal QoS integration during handover. In this view protocols natively targeting handover improvement are extended opening possibilities for new research challenges.

Figure 2.5 depicts the Daidalos goal for a 4G mobility architecture. The architecture considers the separation of local and global mobility. Under this concept, the network is divided into several local domains connected via a core network where mobility is supported through a global mobility protocol (GMP), such as MIPv6 [20] or HIP[26]. Inside a local domain terminal mobility is handled through local protocol operations transparent to the core network, like NeTLMM [28], which are independent of the global mobility protocol. The GMP only operates in case of terminal mobility across local domains.

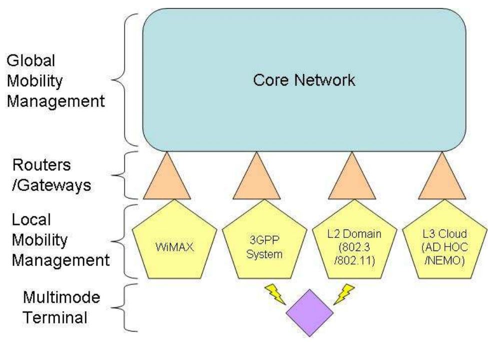

Figure 2.5: 4G Architecture

The design principles of this architecture are built upon local and global mobility partitioning and include the following advantages:

- Making LMP and GMP independent of each other allows operators to choose their preferred LMP or even avoid the maintenance of a mobility infrastructure by direct use of GMP for mobility inside their networks;

- Performance is improved via reduced signalling outside the local domain, and between the terminal and the access network (e.g. reduced signalling over the air and in the core);

- Support for terminals that only operate with global mobility protocols (e.g. HIP or MIPv6 installed);

- Seamless handover support for terminals with enhanced mobility functions;

- Reduced overhead for data path in the air interface;
- The partitioning of LMM and GMM helps in security related issues;
- Architecture of the terminals with enhanced mobility functions common to all LMPs and independent of their particularities. Standardization of these functions and associated signalling between the terminal and the access network;
- Interoperation with 3GPP and WiMAX networks which are seen as local domains from the global domain perspective.

The terminal operations resulting from the above principles are further detailed in the following:

- Terminals with only global mobility support implement only the basic GMP with no enhancements. They are provided with transparent mobility inside local domains and can move between local domains using the GMP. Due to the lack of enhanced functions, handovers for this type of terminals are not optimized;
- Terminals with enhanced mobility functions implement the GMP plus some additional functionality for optimized operation, including seamless handovers. The enhanced mobility functions implemented by the terminal need to be standardized and are independent of the LMPs.

As result, LMDs can potentially implement network based mobility management schemes for resource control thus efficiently combining handover procedures with user allocation. Hence, dividing mobility control in GMD and LMD enables network controlled handovers, regarded as an essential add on when targeting heterogeneous wireless access.

Going further, there are a number of possible choices that the above architecture gives to mobile operators:

- An operator has the flexibility to choose any LMP to handle mobility in its own network. Alternatively, a mobile operator may decide to directly use the GMP to support mobility in its own network and thus avoid installing any mobility related infrastructure. In this latter case mobility functions are supported by equipment located in other networks outside the operator's domain;
- Running its own LMP brings a number of advantages for the operator. On one hand, performance is improved as a result of avoiding signalling outside the local domain. Furthermore, running a LMP allows for more optimized and efficient operations (like e.g. handovers), as mobility is handled locally and closer to the terminal. Finally, from an operational viewpoint LMPs give the operator a greater control of mobility operations as these do not depend on functions external to the operator's network;
- One of the key advantages of the architecture is that it nicely integrates legacy networks that are running their own mobility management scheme, such as the 3GPP Long Term Evolution Architecture and the WiMAX Forum mobile architecture. Indeed, a terminal that supports the functionality described above plus 3GPP/WiMAX would be able to interoperate in this architecture, with 3GPP/WiMAX acting as local domains, without requiring any further modification.

With the above architecture, interconnecting different access operators becomes simpler, as each one can not only manage his network at will, but even select which technology is best for him, while retaining the overall interoperation. In addition, this architecture also relieves the requirements on the terminal side since as long as the operator is running a LMP, it can provide mobility transparently within the local domain to terminals that do not implement any mobility function.

2.5 Requirements

From this chapter it is possible to infer a series of requirements for the support of all these concepts, particularly when integrated into a single 4G architecture. In a broad sense, these requirements can be:

- Support of multiple technologies in an abstract way, allowing fully integrated heterogeneous environments;
- Supporting mobility schemes with the objective of maintaining a user session throughout a call;
- Realizing the existence of network management entities responsible for maintaining system integrity, as well as providing valuable and reliable services to users;
- Enhancing information existing in the network (used to take actions and make decisions regarding load balancing, call admission, etc.) with information from the terminals, allowing the network to have indications from the terminal's view of the network;
- Division of the network into different hierarchical domains, separating the access functions that deal directly with individual terminals, from the core mechanisms which aim to maintain the network's operability;
- Consider Quality of Service for costumers enabling an enjoyable experience while using the services of the network.

These requirements are the motivation to study the IEEE 802.21 standard as an enabler for the integration of these concepts. In the next chapter a study of the standard's services and mechanisms is presented. Chapter 4 presents the study of an existing simulation model where a terminal is enabled with 802.21 mechanisms. Chapter 5 presents simulation results obtained from implementing network and terminal extensions to the previous simulation model, in order to enhance it with new code able to support Network-Initiated Handovers. This standard was also studied for application to 4G architectures, where a general overview is presented in chapter 6.

Chapter 3

The IEEE 802.21 Media Independent Handovers Standard

This chapter provides a brief overview of the IEEE 802.21 standard [1], highlighting where possible information relevant for the development of the thesis.

3.1 Definition

802.21 is a standard from the IEEE, currently under development, whose purpose is to define extensible media access mechanisms that may facilitate handovers between 802-based protocols and cellular systems, as well as optimizing handovers between 802 systems, whether they are wireless or not. More specifically, it aims to provide link layer intelligence and other related network information to upper layers, or to the mobility management entity responsible for handover decision making, to optimize handovers between heterogeneous media.

The standard supplies a framework allowing for transparent service continuity while a mobile node is switching between heterogeneous technologies. For session continuity it is important that the framework can properly identify the mobility-management protocol stack residing in the network elements supporting the handover.

Both hosts and network entities will feature a new entity, the Media Independent Handover Function (MIHF), available within their protocol stack, with correspondent Service Access Points (SAPs) and associated primitives, to access the services therein.

The services available are:

- **Media Independent Event Service** Provides event classification, filtering and reporting, corresponding to dynamic changes in link characteristics, status and quality;
- **Media Independent Command Service** Enables MIH clients to manage and control link behavior related to handovers and mobility;
- **Media Independent Information Service** Provides an information repository, for query and response, supplying details on the characteristics and services provided by the serving and surrounding networks.

The information exchange occurs between lower layers and higher layers, taking always as a reference the MIH Function. Furthermore, the information can be shared locally, within the same protocol stack, or remotely, between peer MIHFs.

Primitives are used to access these services. A primitive is a unit of information which is sent from one layer to another. There are four classes of primitives: Request, Confirm, Indication and Response. The request is issued by the layer that wants to get the services or the information from another layer, and the confirm is the acknowledgment of the request. The indication is the notification of the information to the layer that requested the service, and the response is the acknowledgment of the indication.

A primitive consists of five fields: the protocol layer identifier to which this primitive should be sent, the protocol identifier through which protocol entity this primitive should be sent, the primitive class (i.e., request, confirm, indication, or response), the primitive name, and parameters.

There are three different usages of primitives:

- **To immediately provide L2 information to upper layers** A "Request" primitive is an acquisition request for L2 information. As a "Confirm" primitive, L2 information returns immediately.

- **To notify upper layers of L2 events asynchronously** "Request" and "Confirm" primitives are used just for registration. When an event occurs, an "Indication" primitive is asynchronously delivered to an upper layer.

- **To control L2 actions from upper layers** A "Request" primitive is a request for operation. Ack or nack returns immediately as a "Confirm" primitive.

The primitives' purposes are to collect link information and control link behaviour during handovers.

3.2 Handovers

The 802.21 specification defines mechanisms to support MN-initiated, MN-controlled, network-initiated and network-controlled handovers. It also considers handovers for mobile users and stationary users. In the first case a handover can occur due to changes in wireless link conditions, or due to gaps in wireless coverage that result from mobile node movement. In the second case, the environment surrounding the stationary user can change, making one network more preferable than another. Another example can be the selection of an application which requires a transition to a higher data rate link.

The 802.21 standard aims to facilitate handovers of different methods, classified as "hard" or "soft" depending on whether they are made break-before-make or make-before-break.

Handovers can be affected by diverse factors such as service continuity, application class, etc. The standard supplies essential elements which address the following handover factors:

- **Service Continuity:** this refers to the continuation of the service during and after the handover, minimizing data loss and break time without requiring user intervention, for the cases where the handover is done between two different access networks or between two different points of attachment of a single network. The standard can, for instance, supply information to an application regarding the QoS available on a candidate network, and determine if a handover is viable of not, considering if the required QoS is supported.

- **Application Class:** the 802.21 standard can provide characteristics for application aware handover decisions, taking into consideration the different tolerance for delay and data loss.

- **QoS:** this is a very important factor to consider in handover decision making. In general, the decision is taken by considering the network which best provides the most appropriate level of QoS. This standard specifies the means by which QoS information can be obtained for each of the supported access networks and be made available to upper layers involved in the process of handover decision making.

- **Network Discovery and Selection:** the standard defines the network information to aid in providing new possibilities for network discovery and selection, as well as defining the means to obtain it.

- **Security:** the message exchange between the MIH services of the MN and a Point of Attachment (PoA) must be secured until the MN has a secure connection with the PoA. This association can either be achieved through lower or higher layer security mechanisms. The 802.21 standard supports the means for security information to be made available to upper layers for setting up secure connections.

- **Power Management:** the 802.21 standard allows the availability of media dependant information to higher layers and remote entities, without having to use the specific media to obtain it. This allows for efficient "sleep" or "idle" modes that optimize the use of battery.

- **Mobile Node Movement:** the 802.21 standard allows for the supply of fresh link information to local and remote entities for efficient handover decision making.

The 802.21 standard then proposes to enhance handover operation through cooperative use of both the mobile node and the network infrastructure, allowing for access to handover related information such as neighbourhood cell lists, location of mobile devices, high-level services availability, etc.. The aim is to optimize network selection enabling better and increased handover opportunities. Both the mobile node and the network can make network selection decisions, contributing to an optimum network selection, based on the information available.

One has to note that handover control, policies and algorithms are handled by communication system elements which fall away from the scope of the standard. Nevertheless, some aspects of the overall handover procedure must be described in order to better clarify the role of the MIH function services.

3.3 The Media Independent Handover Function

The Media Independent Handover Function is a logical layer in the mobility management protocol stack, both in the mobile node and the network elements. Its purpose is to aid and facilitate handover decision making though the supply of inputs and context to the upper layers, for handover decision and link selection. Figure 3.1 displays the MIHF regarding its placement between Upper Layers and Lower Layers. One of the key goals of the MIHF is to facilitate the recognition of handover occurrence, as well as the discovery of information that enhances effective handover decisions.

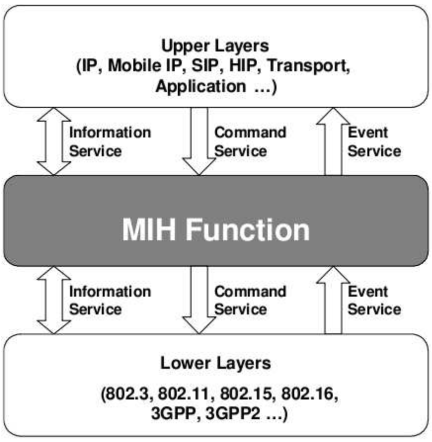

Figure 3.1: MIH Function and offered services

The MIHF provides abstracted services to higher layers, offering a single unified interface to them, through technology independent primitives. The MIHF itself relies on technology specific interfaces to communicate with the lower layers. The specification of these interfaces is not in the scope of this standard, since they are already defined as Service Access Points within their respective standards.

3.3.1 The Communication Model

MIH Functions in different entities can communicate with each other, for instance, to exchange information about the network with the terminal, in order to aid in the handover decision making process. The standard defines the communication relationships between the elements of the network that have a MIHF.

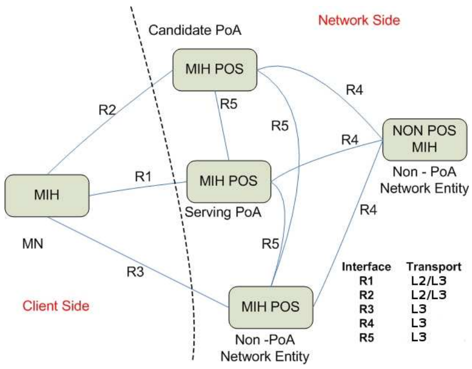

Figure 3.2: IEEE 802.21 reference model

The IEEE 802.21 communication reference model (figure 3.2) specifies interfaces between mobile devices and points of attachment to the network and between network nodes in the network. The information exchange occurs between lower layers and higher layers, taking always as a reference the MIH Function. Furthermore, the information can be shared locally, within the same protocol stack, or remotely, between different network entities. As shown in figure 3.2 interfaces R1 and R2 are specified at layer two and layer three, while interfaces R3, R4 and R5 are specified only at layer three aiming at technology independence, thus allowing approaches such as NIHO. R3 is a special case where layer two communication is possible in specific cases, like Ethernet bridging, or multi-protocol label switching (MPLS).

The Point of Attachment (PoA) is an endpoint of the network side that includes a MN as the other endpoint. This is associated to an interface instead of the whole node.

The Point of Service (PoS) entity is an element in the network side whose MIH Function exchanges messages with a MN's MIH. One has to note that a network entity can supply more than one Point of Service and provide different combinations of MIH services to a MN, based on subscription or roaming conditions. Also, a network entity composed of more than one interface has the PoS associated to the network entity itself and not with just one of its interfaces.

In case of a network model that includes a MIH proxy, the MIH network entity that communicates with that proxy does not have a direct communication with the MN and thus is not a PoS for the MN. Nevertheless, the same network entity may still act as a PoS for a different MN.

Between the different instances of MIH Functions, the communication can be made as following:

- MIH on MN and MIH PoS on the serving PoA, where L2 or L3 communication can occur;
- MIH on MN and MIH PoS on a candidate PoA where L2 and L3 communication can occur;
- MIH on MN and the MIH PoS on a non-PoA entity where L3 communication can occur, and also L2 is possible through Ethernet bridging, MPLS, etc.
- MIH on the PoS and a non-PoS in another network entity can encompass L3 communication;
- MIH on two PoS instances in distinct network entities can encompass L3 communication.

3.3.2 The MIH Services

The MIH Function provides synchronous and asynchronous services through well defined Service access points for link layers and MIH users. These services have the purpose to manage, determine and control the state of the underlying interfaces.

3.3.2.1 Media Independent Event Service

Mobile nodes using MIH services receive indications from link layers for asynchronous operations like the Event service. Events can indicate changes in state and transmission behaviour of the physical, data link and logical link layers, and also, with a pre-indicated confidence level, predict changes in these layers.

Events can have origin in the MIH Function, called MIH Events, or from the lower layer, called Link Events. The typical flow of events has its origin at lower layers and are sent to the MIH Function. Upon registration to the MIHF, a high layer entity can receive the events indicated in the registry message.

Events can also be local or remote. Local events are propagated across layers within the local stack of a single device, whereas remote events traverse across the network from one MIH function to a peer MIH function. These events can then traverse from the local MIH function to the local upper layer entity, supposing prior registration to receive that event.

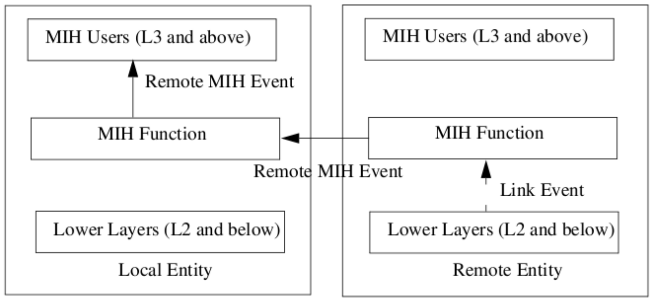

Figure 3.3: MIH Events

Event registration is a mechanism that allows upper layers to indicate to the MIH which events they are interested in receiving. The event service typically is used to aid in detecting the need for handover. For example, it can supply an indication to the higher layers that the link will go down in a near future (through analysis of the signal level and crossing a threshold) which can be used to prepare a new point of attachment ahead of the current one going down.

From the recipient's point of view, these events are mostly advisory in nature and not mandatory. Also, higher layer entities might have to deal with freshness, reliability and robustness issues related to these events, specially in the case that they are remote.

The Media Independent Event Service supports several types of events:

- **MAC and PHY State Change events:** which correspond to changes in MAC and PHY state;
- **Link Parameter events:** triggered due to change in link layer parameters. Can be triggered synchronously (i.e. regularly) or asynchronously (i.e. when a threshold is crossed);
- **Predictive events:** express the likelihood of a future change in a certain property, with a certain degree of certainty, based on past and present events;
- **Link Synchronous events:** give indications of precise timing of L2 handover events that are useful to upper layer mobility management protocols;
- **Link Transmission events:** indicate the transmission status of higher layer PDU's by the link layer. This information can be used by upper layers to improve buffer management for achieving low-loss or loss-less handovers.

Currently defined events are:

- **Link Up** This notification is delivered from the lower layers of a terminal or a PoA, to an MIH-user (local or remote) when a layer 2 connection is successfully established with mobile node;
- **Link Down** This notification is delivered from the lower layers of a terminal or a PoA, to an MIH-user (local or remote) when a layer 2 connection is disconnected due to a certain reason with a mobile node;
- **Link Going Down** This message is transmitted when a layer 2 connectivity is expected (predicted) to go down (Link_Down) within a certain time interval;
- **Link Detected** This message is transmitted when a new link has been detected;

- **Link Parameters Report** This message indicates changes in link parameters that have crossed pre-configured threshold levels;

- **Link Event Rollback** This message is used in conjunction with Link_Going_Down. If the link is no longer expected to go down in the specified time interval, a Link_Event_Rollback message is sent.

- **Link Handover Imminent** This message indicates that a link layer handover decision has been made and its execution is imminent;

- **Link Handover Complete** This message indicates that a link layer handover has been completed.

When exchanging events and commands between hosts, extreme caution needs to be exercised in establishing communications securely. Unless endpoints are identified and authorized for communications, one or more service entities may be affected by state changes as a result of inappropriate registration, event or command reception. [37] identifies that unauthorized third parties impersonating legitimate ES entities can generate hints, attackers can cause damage to communications to or from a mobile node. Event indications are therefore likely to be interpreted as hints by other protocol layers, without necessarily modifying protocol state. Unless an appropriate security mechanism and a security model are adopted, attacks may be particularly simple on a wireless medium, where even valid event messages may be replayed at a later stage by an attacker. Command services experience most of the same attacks as event services, although their effects may be even more severe, since commands imply state changes.

3.3.2.2 Media Independent Command Service

The Media Independent Command Service allows higher layers to configure, control and get information from the lower layers. The information provided by these commands is dynamic in nature comprising of link parameters such as signal strength, link speed, etc. These commands can be issued by higher layers, called MIH commands, as well by the MIH Function, called Link Commands. Typically, messages propagate from the upper layers to the MIH Function, and then from this to the lower layers.

MIH Commands may be local or remote. Local MIH Commands are sent by Upper Layers to the MIH Function in the local stack, whereas remote MIH Commands are sent by Upper layers to the MIH Function in the peer stack.

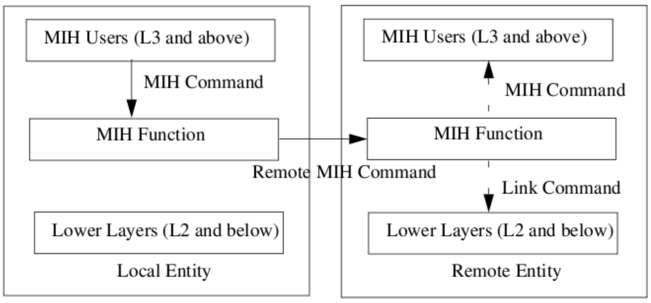

Figure 3.4: MIH Commands

The commands generally carry the upper layer decisions to the lower layers on a local or remote device entity. An example is a policy engine entity to request a mobile node to switch link.

Currently available commands are:

- **MIH Get Status** This command is used to discover the status of currently available links;
- **MIH Switch** used to switch an active session from one link to another;
- **MIH Configure** used to configure the link and control the link behavior;
- **MIH Configure Thresholds** used to configure link parameter thresholds;
- **MIH Scan** used to scan a list of PoAs for a specific link type;
- **MIH Handover Initiate** used to communicate an intent of handover initiation. The handover may be initiated either by the client or the network.
- **MIH Handover Prepare** used to prepare the new link resource for impending handover and to query the available link resource of the candidate network.
- **MIH Handover Commit** used to communicate the intent to commit to a handover request to a specific link and PoA.
- **MIH Handover Complete** used to communicate the status of handover operation. Also used to communicate the preferred action to be taken w.r.t resources associated with the previous connection. If the handover is successful, the resources may be released;
- **MIH Network Address Information** used to discover a mobile node's network address related information before handover. Upon reception of this message, the receiving remote MIHF may interact with its upper layers to obtain the information or forward this message to another MIHF which may have the requested information. Serving FA Address / Serving Access Router Address/ DHCP Server Address parameter is chosen depending on what kind of IP address configuration methods is set. For example, if the IP Address configuration method is IPv4 dynamic configuration (DHCPv4), then the IP address included in "Serving FA Address / Serving Access Router Address/ DHCP Server Address" is the DHCP Server address (IPv4).
- **MIH Get Information** used to retrieve the Information Elements provided by the information service.

3.3.2.3 Media Independent Information Service

The Media Independent Information Service provides a framework and corresponding mechanisms by which a MIHF entity can discover and obtain network information, with the purpose to facilitate handovers.

This service provides a set of Information Elements, the structure to store them and its representation, and a query/response mechanism for transferring the information. Different types of Information Elements may be necessary depending on the type of handover. For example, for horizontal handovers across different PoAs of the same access network, information from lower link layers can be sufficient. But, in the case of vertical handovers, the mobile node may move across different access networks. Then, it is necessary to select an appropriate PoA in the new network based in good lower link connectivity as well as in the availability of higher layer services.

This information can be made available via both lower and upper layers, and, if necessary, be obtained through a secure port. The information service also provides access to static information such as neighbour reports, which help in network discovery. It also includes more dynamic information such as channel information, MAC addresses, security information, etc. This, in conjunction with high layer information (such as application requirements) may help in a more effective handover decision. The set of different Information Elements may evolve. Also, there is a need for flexibility and extensibility in the way this service provides the information, since, for example, the list of available access networks is always evolving. For this, the standard defines a representation schema, which allows a client of MIIS to discover the entire set of different access networks and Information Elements supported, in a flexible and efficient manner. This schema can be represented in multiple ways, such as Resource Description Framework (which is based on *eXtensive Markup Language* (XML)), *Abstract Syntax Notation number One* (ASN.1) (which is used in 802 MIBs), variants or a simple *Type Length Value* (TLV) representation of different information elements.

This service allows access to heterogeneous information about networks, to be accessible by both the mobile node and the network. A media-independent neighbour graph may be abstracted through the neighbour reports of specific technologies. This capability allows the terminal to use its current access network technology to query information about other technologies without activating that interface. One can also, for example, obtain knowledge of supported channels by different PoAs without resorting to scanning or beaconing. The MIIS and MICS information could be used in combination by the mobile node, or network, to facilitate the handover.

The information elements can be classified in three groups:

- **General Access Network Information:** provide a general overview of the different networks, such as list of available networks, operators, roaming agreements, cost of connections, security and quality of service;
- **Points of Attachment Information:** such as addressing, location, data rates, etc;
- **Other:** information specific information and vendor specific.

In [38] the deployment of information services is discussed. For these information services, it is possible that network information may be either centrally stored on a server or distributed in each individual access network. Presumably, an L2 or L3 based mechanism to identify or discover a valid information server would be required. In order to accomplish this, it will be necessary to describe IS discovery and specify transport services over IP. Interactions with IP in delivering handover services over IP therefore need consideration in the IETF, both for use with 802.21 and for other instances of handover services.

3.3.2.4 Information Representation

A schema defines structure of information. A schema is used in the 802.21 information service to define the structure of each information element as well as the relationship among different information elements supported. The MIIS schema is classified into two major categories.

- Basic schema that is essential for every MIH to support
- Extended schema that is optional and can be vendor specific

In [39] a new DHCPv4 option is defined for discovering MIIS information.

3.3.3 Service Access Points

Service Access Points are used to exchange messages between the MIH Function and other planes, using a set of primitives that specify the information to be exchanged and the format of those information exchanges.

The 802.21 standard includes the definition of media independent SAPs, the MIH_SAP, that allow the MIH Function to provide services to the upper layers. These upper layers are called the MIH Users and use the services provided by the MIH Function through the MIH_SAP. The MIH Users have to register to the MIH Function in order to obtain access the MIH generated events and the Link Events, which are generated by the layers below the MIH but are passed on upwards through the MIH Function. MIH Users may also send commands to the MIH Function. Also, MIH Function entities may also send remote commands to other remote MIH Function entities.

The standard also includes recommendations to define or extend existing media-dependent SAPs. These allow the MIH Function to use services from the lower layers of the mobility-management protocol stack and their management plane. All the Link Events generated at lower layers and all the Link Commands sent through the MIH Function are part of the media specific MAC/PHY SAPs and are already defined by their specific technologies.

Lastly, the MIH_NMS_SAP interfaces with the Network Management System, responsible for configuring parameters in the MIHF.

These concepts can be combined in the following general MIH Reference Model:

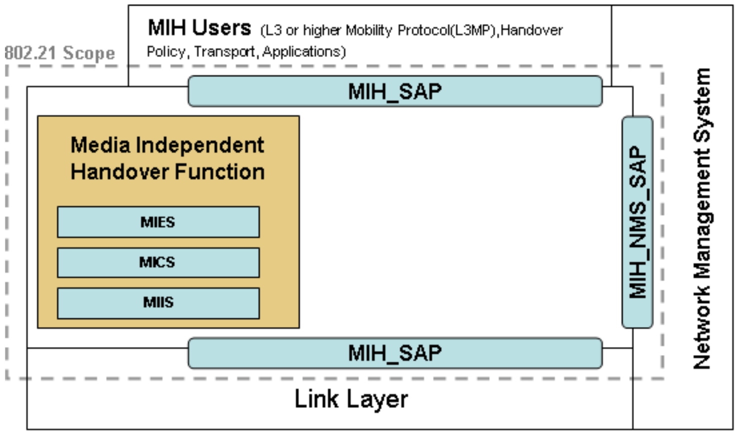

Figure 3.5: MIH Reference Model

3.3.4 The MIH Protocol

The MIH protocol defines the format of the messages (i.e. MIH packet with header and payload) that are exchanged between remote MIH entities and the transport mechanisms that support the delivery of the messages. The selection of the transport mechanism is dependent on the access technology that connects the MN to the network. These messages are based on the primitives which are part of Media Independent Event service, Media Independent Command service and Media Independent Information service.

MIH messages require reliability for remote communication on an end-to-end basis to ensure the receipt of data to the destination. Reliability may be provisioned with an optional acknowledgement service as part of the MIH protocol. The source endpoint may optionally request for an MIH ACK message to ensure successful reception of a certain event, command or an information service message. When this MIH ACK is received by the source, it may conclude that the message was reliably delivered to the destination. In case of a lost MIH ACK message, the source shall timeout and retransmit the same MIH message. This timer may be related to the RTT between the two nodes.

This mechanism is used through two bits in the MIH message header: the ACK-Req bit is set by the source and the ACK-Rsp is set by the destination. The underlying transport layer takes care of verifying the MIH message integrity, and thus not required at MIH Function level.

The packet payload for these services may be carried over L2 management frames, L2 data frames or other higher layer protocols.

3.3.4.1 Services Provided by the Protocol

The MIH Protocol provides the following services:

- **MIH capability discovery:** so a MIH Function can discover MIH Functions on other entities. This allows for the negotiation and selection of an optimum transport for communication, as well as discovering the list of supported events and commands.
- **MIH remote registration:** so that a MIHF in an entity can receive remote events;
- **MIH message exchange:** using the messages supplied by the MIH services.

3.3.4.2 Protocol Identifiers

Successful communication between two MIHF peers requires addressing and identifying the session between them. Three identifiers exist to serve this purpose.

- **MIHF ID** Media Independent Handover Function Identifier is an identifier that is required to uniquely identify MIHF end points for delivering the MIH services that are subscribed for it. As described in earlier clauses, MIH registration enables two MIH entities to create an MIH pairing based on the policies configured in the MN or after capability discovery mechanism. MIH network entities may not be able to provide any MIH services to the MN without this signaling from MN. Use of MIHF ID also enables the MIH protocol to be transport agnostic. MIHF ID may be assigned to the MIHF during configuration process. MIHF ID is used during MIH registration and all messages that are required to identify the end points that are currently not in a session;
- **Session ID** Session Identifier is assigned during a session establishment for remote registration. Each remote registration needs to be associated with the MIHF that keeps the registration information for a particular MIH service;
- **Transaction ID** A MIHF Transaction is a combination of an MIHF Request or Indication message and the corresponding MIHF Response message (if applicable) that are exchanged between two MIHF peers. Transaction Identifier is an identifier that is used with every MIH initiator and its response message. This is also required to match each request, response or indication message and its acknowledgment. This Identifier shall be created at the node initiating the transaction and it is carried over within the fixed header part of the MIHF frame. In order to handle the case of duplicate transaction identifiers at the responding node, this identifier is used in conjunction with the Session Identifier. In messages where Transaction ID is not used it is set to 0.

3.3.4.3 Frame Format

The frame format for MIHF has the following appearance:

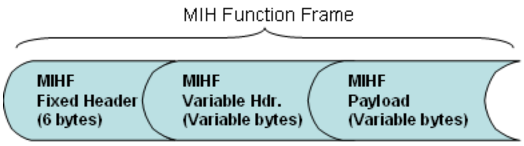

Figure 3.6: MIH Function Frame Format

The MIH Function Header consists of two parts: one with fixed length and another with variable length. The first one accommodates the essential information which is present in every packet and is used for parsing and quick analysis. The second carries information which is optional in some cases.

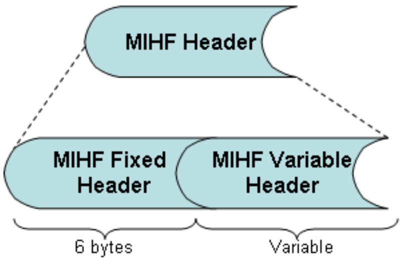

Figure 3.7: MIH Function Header

Table 3.1 lists the fixed header fields.

Field Name	Size	Description
Version	4	This field is used to specify the version of protocol used.
ACK-Req	1	This field is used for requesting an acknowledgement for the message.
ACK-Rsp	1	This field is used for responding to the request for an acknowledgement for the message.
Reserved	4	This field is intentionally kept reserved. In un-used case, it all the bits of this field are to be set to '0'.
MIH Message ID - Service Identifier (SID) - Operation Code (Opcode) - Action Identifier (AID)	16 4 3 9	Combination of the following 3 fields. Identifies the different MIH services, possible values are: 1:System Management 2:Event Service 3:Command Service 4:Information Service Type of operation to be performed with respect to the SID, possible values are: 1:Request 2:Response 3:Indication This indicates the action to be taken w.r.t. the SID
Number of additional header Identifiers	8	Indicates the no. of header identifiers (TLV for each) included in the variable MIHF header part
Transaction ID	16	This field is used for matching Request and Response as well as matching Request, Response and Indication to an ACK.
Variable Load Length	16	Indicates the total length of the variable load embedded into the MIH Function frame and is the sum of MIH variable header length and MIHF payload length.

Table 3.1: MIH Function Fixed Header Description

The variable header presents additional elements that help analyzing and coordinating the payload. These identifiers are represented in Header-TLV format.

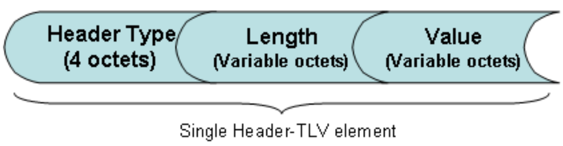

Figure 3.8: Header TLV Format

The Header Type field indicates the type of header identifier embedded in the Value field. The Header Length field indicates the length of the Value field, and the Value field indicates the value of the element.

Table 3.2 lists the function messages.

No	MIH Action Identifier	MIH Opcode	MIH Service ID
1	MIH Capability Discover	Request, Response	System Management
2	MIH Event Register	Request, Response	Event Service
3	MIH Event Deregister	Request, Response	Event Service
4	MIH Link Up	Indication	Event Service
5	MIH Link Down	Indication	Event Service
6	MIH Link Going Down	Indication	Event Service
7	MIH Link Detected	Indication	Event Service
8	MIH Link Parameters Report	Indication	Event Service
9	MIH Link Event Rollback	Indication	Event Service
10	MIH Link Handover Imminent	Indication	Event Service
11	MIH Link Handover Complete	Indication	Event Service
12	MIH Get Status	Request, Response	Command Service
13	MIH Switch	Request, Response	Command Service
14	MIH Configure	Request, Response	Command Service
15	MIH Configure Thresholds	Request, Response	Command Service
16	MIH Scan	Request, Response	Command Service
17	MIH Handover Initiate	Request, Response	Command Service
18	MIH Handover Prepare	Request, Response	Command Service
19	MIH Handover Commit	Request, Response	Command Service
20	MIH Handover Complete	Request, Response	Command Service
21	MIH Network Address Information	Request, Response	Command Service
22	MIH Get Information	Request, Response	Information Service

Table 3.2: MIH Function Messages

The parameters in MIH Protocol Messages are expressed through TLV encodings.

The messages are divided in messages for System Management Service, for Event Service, for Command Service and for Information Service Category.

3.3.4.4 Capability Discovery

The MIH Function in the terminal, or in the network, may discover which entity in the network, or terminal, supports MIH capability by using the MIH capability discovery procedure. This procedure consists of a handshake and capability advertisement, and is achievable by exchanging MIH messages.

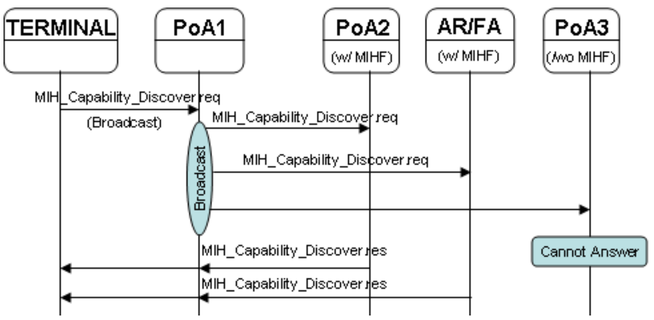

Figure 3.9: MIH capability Discovery Procedure

The mechanism depicted here is directly invoked by the MIH Service. A node that, for example, has just powered up broadcasts a MIH_Capability_Discovery.request message that is further propagated through the access network. Entities that are MIH-enable are able to recognize the message and reply their capabilities with a MIH_Capability_Discovery.response. The first node then acquires information on which nodes around it have 802.21 support, and which events and commands are available.

In [40] it is discussed the possibility of integrating the capability discovery in the common protocol functionality being defined in IEEE 802.21 and IETF. In this case, the MIH service would simply request the delivery of a message from the common protocol functionality, and leave the discovery and resolution procedures to the lower layers. Regarding candidates for integration, we can think of DHCP, SLP and mDNS. More information on IETF involvement is discussed in section 3.5.

3.3.4.5 Encapsulation

Regarding encapsulation, although the standard supports high-level transport protocols such as TCP and UDP, [41] suggests that the MIH message shall be inserted inside a UDP datagram which can fit in either an IPv4 or IPv6 packet. Nevertheless, the draft, for IPv4, does not mention problems such as firewalls or NAT traversal.

MNs will feature a MIH Application that sends and receives MIH messages through UDP on a unique port number that shall be registered and obtained from IANA. It is also assumed that the Mobility Management Entity, in case it exists and it is external to the MN, will have it's IP address discovered as per DHCP special option for discovering IEEE802.21.

The process of a MIH message received in an MN has the following steps:

1. The Network layer receives an IP packet from its lower layers and strips off the IP header, processes it and then forwards it to the appropriate Transport protocol.

2. The Transport layer then receives a UDP datagram. Its headers are in turn removed and processed. The UDP protocol then forwards the contents of its data field to the appropriate Application layer. This is determined by the value of the destination port number. The MIH Application shall have a newly defined port number. Therefore the MIH message would be forwarded to the MIH Application.

3. The MIH Application would then decode the MIH message according to the IEEE 802.21 specifications and shall then react as required.

The steps taken by the MN to transmit a MIH message are symmetric to the steps explained above and the flow shall be in the reverse path as follows:

1. The MIH Application shall generate an MIH message and pass it to the Transport Layer through the newly defined port.

2. The UDP shall encapsulate the data in a UDP datagram and shall set the header fields accordingly.

3. The datagram is then sent to the Network Layer where it is in turn encapsulated in an IP packet and all the header fields of the packet are set accordingly. This packet is then sent to the appropriate lower layer for transmission to the network.

These processes are similar in network nodes also.

3.3.4.6 Reliability

[41] refers to the optional ACK mechanism present in the IEEE 802.21 specification. It proposes timers as a solution for lost or delayed MIH messages. Because the contents of certain MIH messages are more sensitive to delay than others, the values of the timers should be different for the three MIH message types. For example, messages that contain information can be sent periodically to update the mobile node and can have the longest timer. On the other hand, in a network controlled handover scenario for example, the MM may issue a command to a mobile node to handover to a target access technology. Since this node manages the available network resources, such a message would be required to arrive as fast as possible. Thus, the timer associated with command messages should be shorter than those of messages with information. Thus, three timers should be used depending on the type of MIH message that is sent:

- Information timer that is set after the transmission of a message that is related to Information Elements.
- Event timer that is set after the transmission of a message that is related to Events.
- Command timer that is set after the transmission of a message that is related to Commands.

In addition, relying on the service layer to handle all reliability issues opens the question of whether timer values should be based on message transfer latency or application processing latency, and these two can differ significantly.

During the course of this project, several versions of the IEEE802.21 specification where released. After version D1.00 this acknowledgment feature was included.

Reliable delivery for the mobility services may be essential, but it is difficult to trade this off against low latency requirements. It is also quite difficult to design a robust, high performance mechanism that can operate in heterogeneous environments, especially one where the link characteristics can vary quite dramatically. The option of an Acknowledgment mechanism, such as the one used in 802.21, has a number of disadvantages associated with it. The protocol designs ends-up re-inventing a lot of the functionality already available in lower layers at a higher layer where access to information about what is going on in the network is restricted. It also adds to the complexity of the higher layer protocol, and makes successful deployment less certain.

Nevertheless, allowing the option of a reliable transport service means that the additional recovery mechanisms within the service layer can be made very simple and robust because they do not need to be optimised for efficient recovery of message loss within the network.

3.3.5 Specific Technologies Integration

The intent of the IEEE 802.21 standard is to provide generic link layer intelligence independent of the specifics of mobile nodes or radio networks. As such the IEEE 802.21 standard is intended to provide a generic interface between the link layer users in the mobility-management protocol stack and existing media-specific link layers, such as those specified by 3GPP, 3GPP2 and the IEEE 802 family of standards.

The IEEE 802.21 standard defines SAPs and primitives that provide generic link layer intelligence. Individual media specific technologies thereafter need to enhance their media specific SAPs and primitives to satisfy the generic abstractions of the IEEE 802.21 standard. Suitable amendments are required to existing link layer (MAC/PHY) standards of different media specific technologies such as IEEE 802.3, IEEE 802.11, IEEE 802.16, IEEE 802.15, 3GPP, 3GPP2 to satisfy the requirements of generic link layer intelligence identified by IEEE 802.21. However, such specifications of the MIHF interfaces with the lower layers generally do not fall within the scope of this standard. Such interfaces may already be specified as service access points within the standards that pertain to the respective access technologies. This standard may however contain recommendations to amend the existing access technology specific standards when modifications of the lower-layer interfaces may enable or enhance the MIHF functionality.

This section presents the reference models for the 802.16, 802.11 and 3GPP access technologies, being the last two the selected technologies for the simulation work presented in chapters 4 and 5.

3.3.5.1 802.16 Integration

The following figure shows the MIHF for 802.16 based systems.

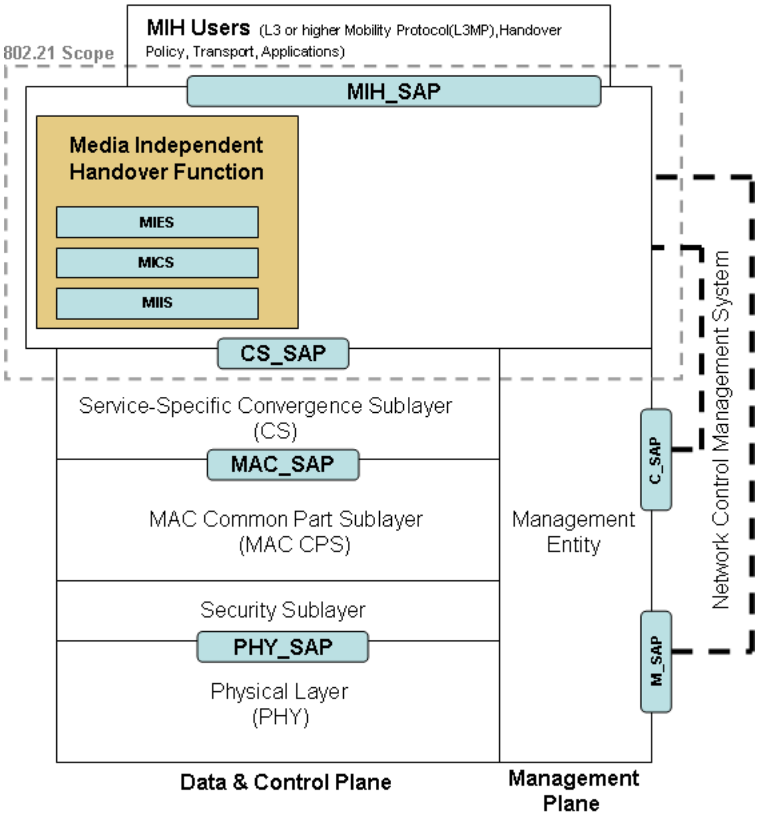

Figure 3.10: MIH Reference Model for 802.61

In the figure it is visible that the M_SAP and C_SAP are common service access points between the MIHF and the NCMS. The SAPs specify the interaction between the MIHF and the control and management entities, also helping in MIH message transportation between different MIHF entities. Also, the C_SAP specifies the interface between the MIHF and the management plane enabling MIH messages to be encapsulated in 802.16 management frames. An important remark is that the primitives supplied by the C_SAP allow messages to be encapsulated before and after a Mobile Node has entered a network with a Base Station, whereas the convergence sublayer (CS) SAP may be used to encapsulate messages after network entry has been accomplished with a BS. This kind of behavior clearly indicates that 802.21 considers the mobile node version of the 802.16 standard, the IEEE802.16e.

3.3.5.2 802.11 Integration

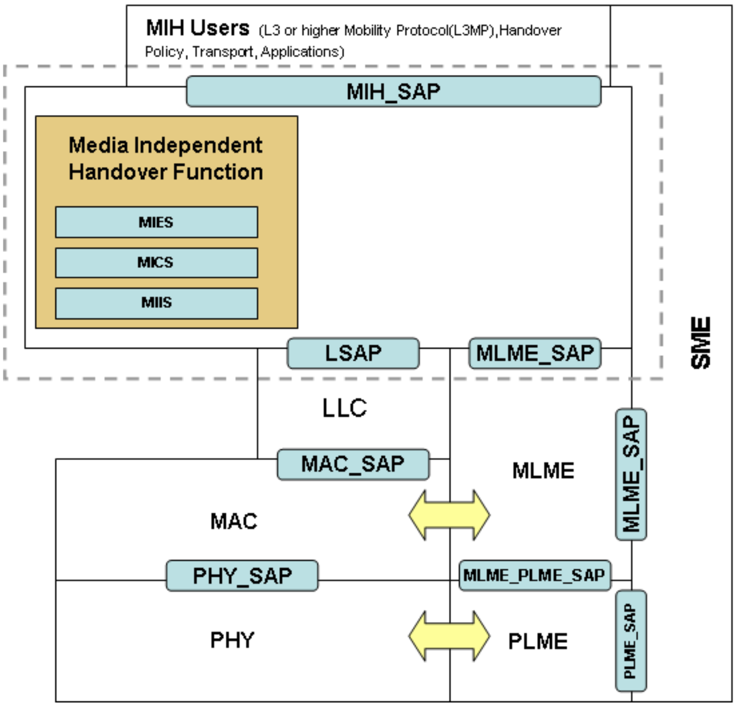

Figure 3.11: MIH Reference Model for 802.11

Figure 3.11 shows the MIH functions for 802.11 stations and network PoA (APs). The LSAP defines the MIH interface to the data plane and may encapsulate MIH payload in data packets. However, since 802.11 does not currently support Class 1 data frames, traffic may be sent over the data plane only when the client has associated with the AP.

The MLME_SAP specifies the interface between MIH and the management plane (MLME) and allows MIH payload to be encapsulated in management frames (such as action frames). Thus primitives specified by MLME_SAP may be used to transfer packets before a station has associated with an AP, whereas the LSAP SAP may be used to transfer packets after association has been established with an AP.

The MIH_SAP specifies the interface of MIH Function with other higher layer entities such as transport, handover policy function and L3 Mobility protocol.

The MLME_SAP includes primitives for system configuration and link state notification/triggers.

3.3.5.3 3GPP Integration

Figure 3.12 illustrates the interaction between MIH Function and 3GPP based system. The MIH Function services are specified by the MIH_3GLINK_SAP. However no new primitives or protocols need to be defined in 3GPP specification for accessing these services. The MIH Function services may be easily mapped to existing 3GPP primitives. The architecture placement of the MIH Function

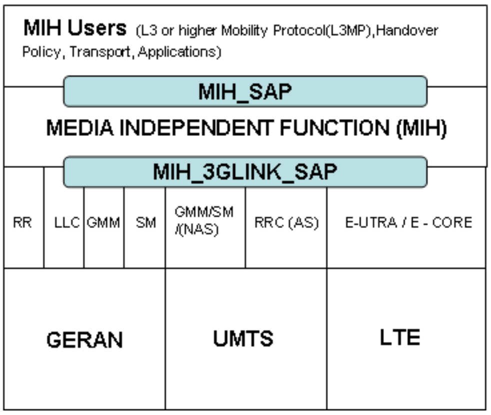

Figure 3.12: MIH Reference Model for 3GPP

shall also be decided by the 3GPP standard. The figure above is for illustrative purposes only and shall not constrain implementations.

For example events received through LAC layers SAP such as "L2.Condition Notification" may be mapped and generated through the MIH-3GLINK-SAP as a Link Up, Link Down or Link Going Down. Likewise events generated at the PPP SAP within the PPP layer, such as LCP-Link-Up or IPCP_LINK_OPEN could be mapped and generated through the MIH-3GLINK-SAP as a Link Up event.

3.3.5.4 Multiple Protocol Stacks

All the previous protocol stacks can co-exist in the same entity. The MIHF is able to determine which media-dependent SAP should be used, via parameters included in the MIH commands invoked by the MIH-users. In the same manner, events originated from a specific link layer technology are correctly identified to the upper layers.

This abstraction provided by 802.21 allows, for example, to obtain information about specific technologies using other links reducing the need to power-up interfaces, etc.

3.4 Media Independent Neighbor Graphs

Neighbor graphs are a mean to represent MIIS information regarding neighborhood. In the specific case of homogeneous networks, network PoAs are of the same kind whereas for heterogeneous networks they are different. Media Independent Neighbor Graphs, or MING, are then composed of heterogeneous and homogeneous elements. Mobile Nodes may obtain media specific graph information directly from the access networks. In those cases, however, the neighbor graph is limited to the media type the mobile terminal is directly attached. 802.21 media independent neighbor information, on the other hand, may help obtain a global view of the network neighbor graph consisting of multiple media types.

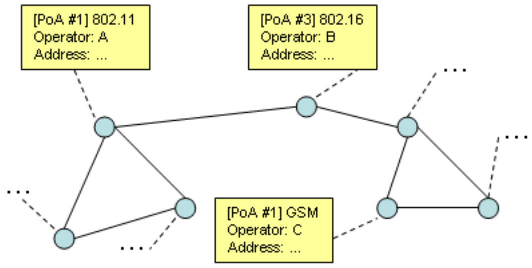

Figure 3.13: Media Independent Neighbor Graphs

The MING contains a set of different types of neighbors relative to a given PoA. This set may be kept by the MIIS functional entity and may eventually help a MN to quickly identify potential candidate PoAs for handover. Other uses exist for MING. For example, the neighbor list supplied by MING can inform the MN about the media types of neighboring networks, in order to avoid powering up unnecessary interfaces, and then scan for neighbor information.

The MING is compiled by the MIIS in the network side. A MING may be manually configured, e.g. manually introducing the neighbors for a given AP, or obtained via the management interface, or other access networks. Rather than incurring such management overhead, the PoA may learn about its neighbors belonging to other media type dynamically through the course of message exchanges. However, how this information is reported and the corresponding responses are outside of the scope of the standard.

A neighboring information element only contains entries of neighboring PoAs that are legitimate neighbors relative to a given PoA satisfying the query. A MING also prevents the addition of fake neighbours. The exact form of implementing the MING is vendor dependent and it is outside the scope of the standard.

If a MN is anticipating a handover, it may request information service from the MIIS entity at the network side. Whenever such request is received, this MIIS entity at the network side replies with a partial set of the MING. This partial concept contains only entries of neighboring PoAs with legitimate relationship of the PoA the MN now makes a connection to. If supported, unsolicited advertisement may also be used for distributing the partial set as well.

3.5 Usage of 802.21 in IETF Standardization Activities

The IETF has been working on how handover services can be used to assist mobility signaling protocols, and has been analyzing interactions between handover services and network/transport layers. At the same time, other groups (e.g. IEEE 802.21) have been taking a more generally applicable view.

The MIPSHOP[1] working group is currently working on common functionality necessary for the support of media independent handover (MIH) protocols, initially focusing on the architectures being developed with the 802.21 working group of the IEEE. Although the MIH protocols carry link layer information, the protocols themselves are not tightly bound to any specific link layer. Instead, they can be carried at the network layer, both within the radio access infrastructure and over the air.

There are on-going activities in the networking community to develop solutions that aid in IP handover mechanisms between heterogeneous wired and wireless access systems including, but not limited to, IEEE802.21. Intelligent access selection, taking into account link layer attributes, requires the delivery of a variety of different information types to the terminal from different sources within the

[1] http://www.ietf.org/html.charters/mipshop-charter.html

network and vice-versa. The protocol requirements for this signalling have both transport and security issues that must be considered. The signalling must not be constrained to specific link types, so there is at least a common component to the signalling problem which is within the scope of the IETF.

A problem statement [42] is in progress, decomposing the overall MIH protocol problem into a common part, and a set of specific MIH services layered over it. The problem statement itself outlines a division of functionality between the common part and the individual services, with the intention that the common part will be generally useful even in non-802.21 architectures. The expectation is that the MIPSHOP working group will develop a solution for the common part which meets the 802.21 requirements as expressed in the problem statement draft.

A second draft [40] outlining more detailed design considerations for the common part has also been prepared which formalises some of the open issues the scope of the common functionality, and lists the key design choices that have to be made and the criteria that should be taken into account in doing so. Although it refers to components of possible solutions as examples, it does not intend to be a solution proposal.

The IETF view on MIH Services is consistent with the architecture defined by 802.21, where the services in question have been identified as the Event, Command and Information Services (ES/C-S/IS). A draft [37] further analyzes the main relevance of ES and CS to IETF and the MIPSHOP WG is the exchange of information related to ES and CS between peer nodes. This is referred to as Remote ES and Remote CS.

3.5.1 Conclusions

This chapter provided an overview of the IEEE 802.21 standard. Here were identified the services, entities and protocol details as defined in the version D1.00 of the draft standard. It should be noted that while this thesis is being written the standard is not yet finalized, potentially making this section outdated in some parts.

Chapter 4

Studying IEEE 802.21 in Mobile Initiated/Controlled Handovers

In this chapter an evaluation of the design of mobile devices implementing the cross layer design advocated by the IEEE 802.21 standard is done. More concretely, this evaluation was done analysing the code of an existing simulation framework implemented in OMNeT++[1] and respective simulation results. In this simulation framework, the study shows how the interaction of lower layers (e.g. WLAN and 3G technologies) with upper layers (e.g. Mobile IP) produces an optimized handover performance. This terminal design for enhanced handovers is the basis for the follow up work presented in the next chapter, where the the simulation framework code was extended to support network and terminal side components for network controlled handovers, allowing for new study and analysis opportunities.

4.1 Motivation

Tomorrow's customers will widely exploit multi-mode terminals (i.e. integrating one or more access technologies) to get better services depending on the environment (e.g. indoor, outdoor). Scenarios in which 3G coverage is complemented by 802.11 or 802.16 access technologies are becoming available. Upcoming standards, such as IEEE 802.21, propose methods to support mobility across heterogeneous technologies.

However, while standards specify functional entities (either implemented in the terminal or in the network) and associated protocol operations, they do not specify configurations of terminal or network components, upon which events are generated.

Considering scenarios in which a terminal can freely move across 3G and WLAN coverage cells, the configuration of the terminal for network detection (e.g. WLAN signal level detection) and attachment is a critical issue. For instance, a user preferring WLAN connectivity (when available) over 3G may need a threshold configuration different from a user preferring 3G. IEEE 802.21 provides a method to configure (upon events or timers) specific thresholds for vertical handovers between 3G and WLAN. However, the values required for a particular scenario are not specified. In this chapter, an evaluation of a simulation framework using a realistic WLAN signal level path loss model ([43] and [44]), where the the effect of terminal speed on handover performance has been investigated, considering a MIH-enabled terminal.

[1] www.omnetpp.org

In the simulation environment, terminals move according to the random way point model at different speeds. Performance of handovers is measured based on Wireless LAN time utilization, packet loss and number of handovers processed. According to the 802.21 standard, the handover algorithm configures the power thresholds, and then handovers are triggered by signals received from lower layers. This study takes into consideration the effect of the mobile terminal speed as a factor for optimal threshold configuration. The results indicate the configuration to be used depending on the value of the primitive *"Link_Configure_thresholds→Link speed"* of the IEEE 802.21 specification [1]. A potential application scenario, could be a IEEE 802.21 based terminal with built-in GPS devices, that could dynamically adjust thresholds' configuration and sampling techniques for WLAN signal level prediction, according with to speed variation.

A number of papers in the literature [45], [46] and [47] have analyzed performance issues of handovers based on Mobile IP between cellular networks. However these works only study the problems related with upper layers (mainly TCP) due to the differences between the two technologies involved. Some previous works (e.g. [48]) study the integration of WLAN hot-spots into 3G networks; these previous works however, are not based on the 802.21 framework in contrast to the one described here. The first work, to the best knowledge of the author, that treats the specific problems of inter-technology handovers based on IEEE 802.21 (taking as Mobility handler SIP) is [49]. However, [49] does not analyze the effect of terminal speed. The WLAN signal level model used in the analyzed simulation framework is based on the model of [43] and [44], where an accurate study on indoor environments is proposed. It is foreseen that indoors WLAN environments complemented by full 3G coverage will be of a great interest in the future.

4.1.1 Terminal Architecture in the Simulation Framework

As seen, the IEEE 802.21 specification defines a middle layer called the Media Independent Handover to centralize functionality related to handover. The studied simulation framework implements the MIH functionality in the OMNeT++ simulation tool. It consists of three elements: the MIH Function, the Service Access Points (SAPs) with their corresponding primitives, and the MIH Function Services (Figure 4.1). The MIH Function (MIHF) is defined in the current IEEE 802.21 specification [1] as

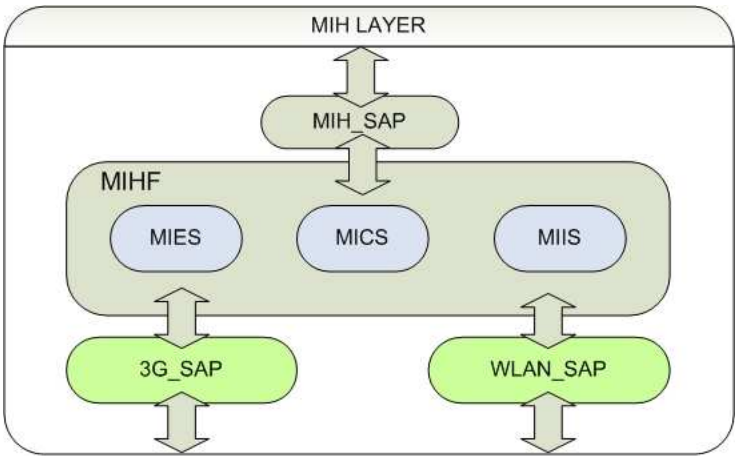

Figure 4.1: MIH Architecture

a logical entity and the specific MIH implementation of the Mobile Node and the network are not included. In fact, it is important to note that in order to facilitate the overall handover procedure, the MIH Function should be defined following a cross-layer design, allowing the communication with the management plane of every layer within the protocol stack.

The Service Access Points (SAPs) are used to enable the communication between the MIH Function and other layers. In the presented architecture there is one technology independent MIH_SAP which allows the communication between the MIH function and upper layers, namely IP, transport, and application. Two technology dependent SAPs are also specified in the simulator: WLAN_SAP and 3G_SAP, which allows communication between the MIH Function and the management plane of the 802.11 link layer and the 3GPP link layer, respectively. Note that every SAP defines certain number of primitives that describe the communication with the services in the MIH Function. Since the implemented scenario does not cover all possible use cases, this implementation only defines the primitives needed for the studied scenario.

The MIH Function is supported by the three basic services specified in the 802.21 draft: events (MIES), commands (MICS) and information (MIIS). As previously stated in section 3, these services can be defined as local or remote. However, since the focus of this simulation implementation relates to specific scenarios where the terminal does not need to discover neighborhood (Information Services) or to receive remote events/commands from the network, only local communication was taken into account.

4.1.2 Modification to the Mobile IPv6 stack

The authors of this simulation framework performed some modifications to the Mobile IP stack of the simulator, in order to maintain handover performance, while enabling 802.21 mechanisms.

Mobile IPv6 [50] signaling (Binding Update BU and Binding Acknowledgement BA) sent by a node for WLAN-3G inter-working, could be lost in the network before reaching the destination or could be lost in the wireless medium when the Mobile Node had poor signal conditions. Taking into account that the signaling is always sent through the new link in this scenario, a signaling loss may occur due to varying WLAN signal conditions when moving from 3G to WLAN. When a BU or BACK is lost the handover at layer 3 is supposed to fail. When the handover fails, the state of the signaling flow can be:

- The BU has not arrived at the Home Agent: the packet flow is reaching the Mobile Node through the old link so no packet loss happens (no handover).
- The BU reaches the Home Agent but the BA is lost: in this case the packet flow starts arriving to the Mobile Node through the new link.

Binding Updates are usually retransmitted upon timeout. If a BA is not received after a timeout expiration, a retransmission is scheduled and the next timeout is set to the double of the original one. This policy is kept until the timeout reaches a maximum (MAX_BINDACK_TIMEOUT is 32 seconds as specified in the Mobile IP RFC [20]).

Since the Mobile Node has no way of knowing if the Binding Update has reached the Home Agent or not after a handover failure, the handover algorithm must proceed with an action to stabilize its state. This action is to perform a handover to the 3G leg. The major modification introduced into the Mobile IP stack is about the way the retransmission of the Binding Updates is handled. The retransmission algorithm as specified in [20] has been omitted and replaced by MIH intelligence which takes the required actions (namely rolling back to the 3G channel) in case BU are lost and timers (about 1.5 seconds) expire.

4.1.3 Handover Algorithm

The handover algorithm implemented in this specific OMNeT++ simulation framework is based on signal thresholds. It relies on the information provided by the Media Dependent layers and the Mobile IP Layer. The handover algorithm reacts upon the reception of three possible signals, which are:

- RSSI (Received Signal Strength Indicator) sample
- Notification about the status of the handover
- Wireless LAN link off message

The handover algorithm is based on two thresholds. The first one, $3G \rightarrow WLAN$ threshold, defines the minimum wireless LAN signal level that must be received in the Mobile Node to trigger a handover from the 3G to the wireless LAN. The second one, $WLAN \rightarrow 3G$ threshold, defines the wireless LAN signal level below which a handover to the 3G leg is triggered.

A handover to 3G can be triggered by two events, when the signal level goes below the $WLAN \rightarrow 3G$ threshold, or when a wireless Link Off message is received.

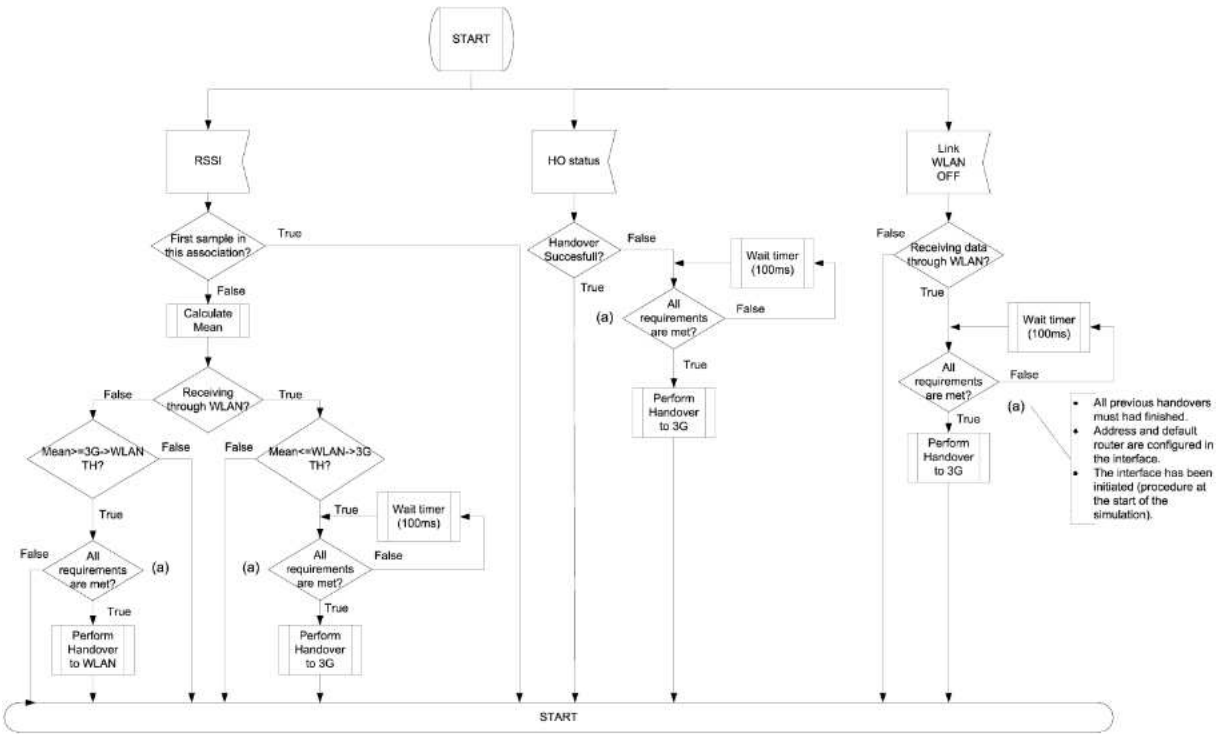

Figure 4.2: Handover Algorithm Flow Diagram

After the MICS (Media Independent Command Service) triggers a handover to the Mobile IP Layer, the handover algorithm is not allowed to perform another handover until the reception of a handover status message informing of the last handover result. If a handover is not successful, the algorithm performs a handover to the 3G part to fix the state of the algorithm. There are different causes for the failure of a handover, for example the BU may not reach the Home Agent or the BU reaches the Home Agent but the Binding Ack is lost.

Before performing a handover some conditions must be satisfied. The interface should be completely configured, with a global routable IPv6 address and default router (Duplicate Address Detection (DAD) procedure completed) associated. Also, all previous handovers should have been completed. If these conditions are not fulfilled the handover is delayed. In the case of handover to WLAN, if the conditions are not met, the handover is skipped until another signal sample arrives. In the case of handover to 3G, the handover is delayed by a timer, waiting for the conditions to be satisfied. The timer has been fixed to 100 ms (default period of the beaconing in WLAN).

4.2 Simulation Modeling

This section describes the simulation modeling implemented to obtain the results from this simulation framework. The handover study was conducted by simulating a Mobile Node attached to the 3G network and performing several handovers between 3G and wireless LAN, varying terminal speeds and round trip time on the 3G link.

The specific analyzed scenario was based on an indoor environment with a wireless LAN cell and full coverage of 3G technology, a scenario that will be a typical deployment in the future. It was noticed that this framework did not cover the WLAN to WLAN handover case. The WLAN to WLAN handover case is extensively studied in [51] which provides a complement the work presented here. Also, this work considers wide space with indoor characteristics (such as an airport) in which the user can move at different speeds.

The simulations considered in this work have the Mobile Node speed varied between 2 m/s and 10 m/s. This value represents an upper limit of the speed expected in a big size indoor scenario. Indeed, all pedestrian speeds are below this threshold.

The movement pattern selected is the Random WayPoint Model. With this model each node moves along a zigzag line from one waypoint to the next one, all the waypoints being uniformly distributed over the movement area.

The data traffic studied is a downstream video, with a packet size of 160 bytes at application layer and inter-arrival packet time of 20 ms (83 kbps)[2]. 60 simulation runs were performed for each experiment. This number was chosen as a trade-off between simulation time and confidence interval size.

4.2.1 WLAN Model

The standard wireless LAN propagation model defined in OMNeT++ is based on free space losses with shadowing and a variable exponential coefficient. The original model implemented in OMNeT++ is suitable for studies that do not analyse in depth the effect of the signal variation. However, the objective of this work relied on the need for a realistic wireless LAN model, suitable for indoor scenarios based on empirical results. For this purpose, the empirical model in [52] was used, which includes variation in the signal due to shadowing and different absorption rates in the materials of the building. The path loss model is the following:

$$
\begin{aligned}
Losses &= 47.3 + 29.4 * log(d) + 2.4 * Y_s \\
&+ 6.1 * X_a * log(d) + 1.3 * Y_s * X_s \\
X_a &= normal(0,1)
\end{aligned}
$$

[2]Notice that usual VoIP codecs can generate bit rates around 80 kbps and therefore their traffic pattern is very similar to the simulated one

$$
\begin{aligned}
Y_s &= normal(-1, 1) \\
X_s &= normal(-1.5, 1.5)
\end{aligned}
\tag{4.1}
$$

Where d is the distance between the Access Point and the Mobile Node.

The power transmitted by the AP and Mobile Node are defined in the UMA specification [53]. According to this specification, the AP transmission power is 15dBm while the Mobile Node transmission power is 10 dBm. Following these specifications, the AP antenna gain is set to 0 dBi while the Mobile Node antenna gain is set to -10dBi. The transmission rate of the wireless LAN is fixed to 11 Mbps.

The OMNeT++ wireless model defines two thresholds, the Sensitivity threshold and the Active Scanning threshold, as can be viewed in figure 4.3. The Sensitivity threshold is the minimum level of signal that the receiver can detect. Real products specifications set this level of signal to -90 dBm[3]. This is the value that was used in the works involving this simulation framework.

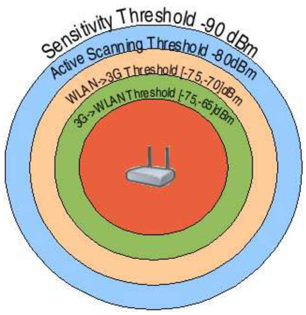

Figure 4.3: Thresholds defined in the simulation environment

The Active Scanning threshold defines the signal level at which the wireless card starts scanning for other APs in order to perform a new WLAN handover. When this level of signal is reached it indicates that the signal level conditions have deteriorated significantly and the Mobile Node detaches from the current AP, starting the scanning procedure for new APs with better signal conditions. The IEEE 802.11b standard does not specify the value for this threshold, its value being design dependant. In the model presented, this value is set to -80 dBm. This value was selected after analyzing via simulations the maximum variability of the wireless LAN signal model. In this work, with this threshold, the Mobile Node will always handover to the 3G link before reaching the sensitivity threshold, in order to avoid connectivity losses. Figure 4.3 also shows the $WLAN \rightarrow 3G$ and $3G \rightarrow WLAN$ thresholds, which represent the signal level that triggers the respective handovers.

4.2.2 3G channel Model

The 3G channel has been modelled as a PPP channel with a connection time of 3.5 seconds, disconnection time of 100 ms, bandwidth of 384 kbps(downlink) and variable delay of 100 to 150 ms per way. These specifications are based on real measurements from a Vodafone data card.

The above PPP channel models the 3G channel when the PDP (Protocol Data Packet) context is activated. The PDP context is a data structure present on cellular networks which contains the subscriber's session information when the subscriber has an active session. When a mobile wants

[3] SMC Networks SMC2532W-B

to use the cellular network, it must first attach and then activate a PDP context. This allocates a PDP context data structure in the cell that the subscriber is currently visiting. The data recorded includes, for example, an IP address. These disconnection and connection times were obtained from measurements in different locations of an office building with a commercial UMTS data card. The round trip time is tuned to typical values of delay in this kind of channel under the same conditions. The connection time is measured as the time elapsed between bringing up the card and the moment when an IP address is assigned to the Mobile Node (activation of a PDP context).

Although the above model takes into account the connection time, in the presented simulations it was assumed that the PDP context is always active, so the value of the connection time does not have any impact. Indeed, these simulations are based on the following two assumptions i) full 3G coverage and ii) 3G link always on, which are realistic assumptions in typical scenarios.

4.3 Evaluation

The handover performance study was done considering the following metrics:

- Wireless Utilization Time
- Number of Handovers
- Packet loss

In a first step, an analysis of the three metrics for different configurations of the thresholds was performed.

Speed varies between 2m/s and 10m/s (Section 4.3.1). In a second step the study was extended with the introduction of the RTT in the 3G channel as additional variable (Section 4.3.2). The way which WLAN signal level is evaluated impacts the overall handover performance, for this reason, several measurement techniques where considered and compared (Section 4.3.3).

4.3.1 Effect of the speed in the thresholds configuration

Figures 4.4 to 4.6 show how the previously mentioned metrics are affected by the different mobile terminal speeds and different $WLAN \rightarrow 3G$ threshold values.

Figures 4.4 and 4.5 respectively show the amount of time the Mobile Node is connected to the WLAN and the number of handovers performed by the Mobile Node.

It can be observed that while the number of handovers decreases when more stringent thresholds are configured, the Wireless utilization time increases. This shows that the proposed algorithm (based on two thresholds), if properly configured, can optimize the wireless utilization time by reducing the number of useless handovers. Another interesting observation is that as the speed decreases, the difference in the wireless utilization time for different speed values increases.

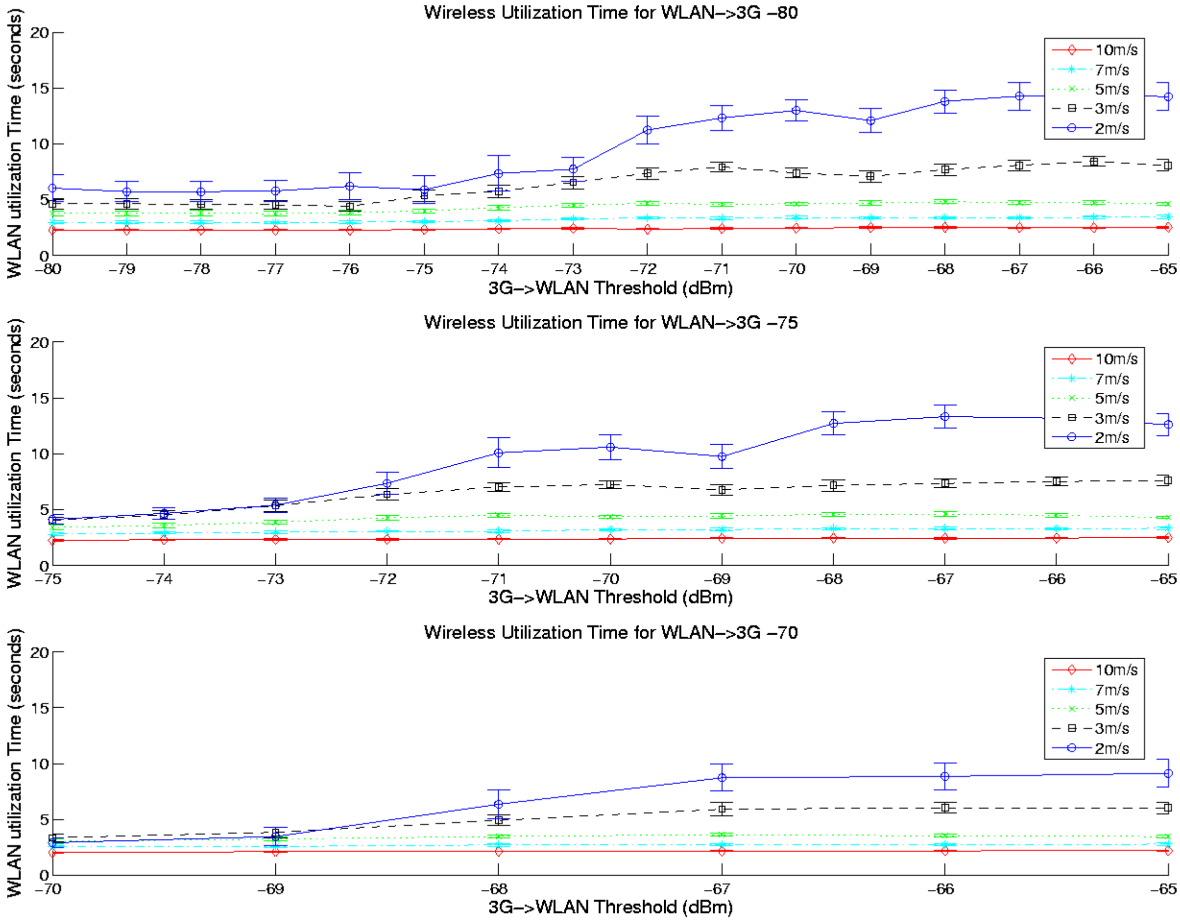

Figure 4.4: Wireless Utilization Time for several speeds (RTT 3G 300ms)

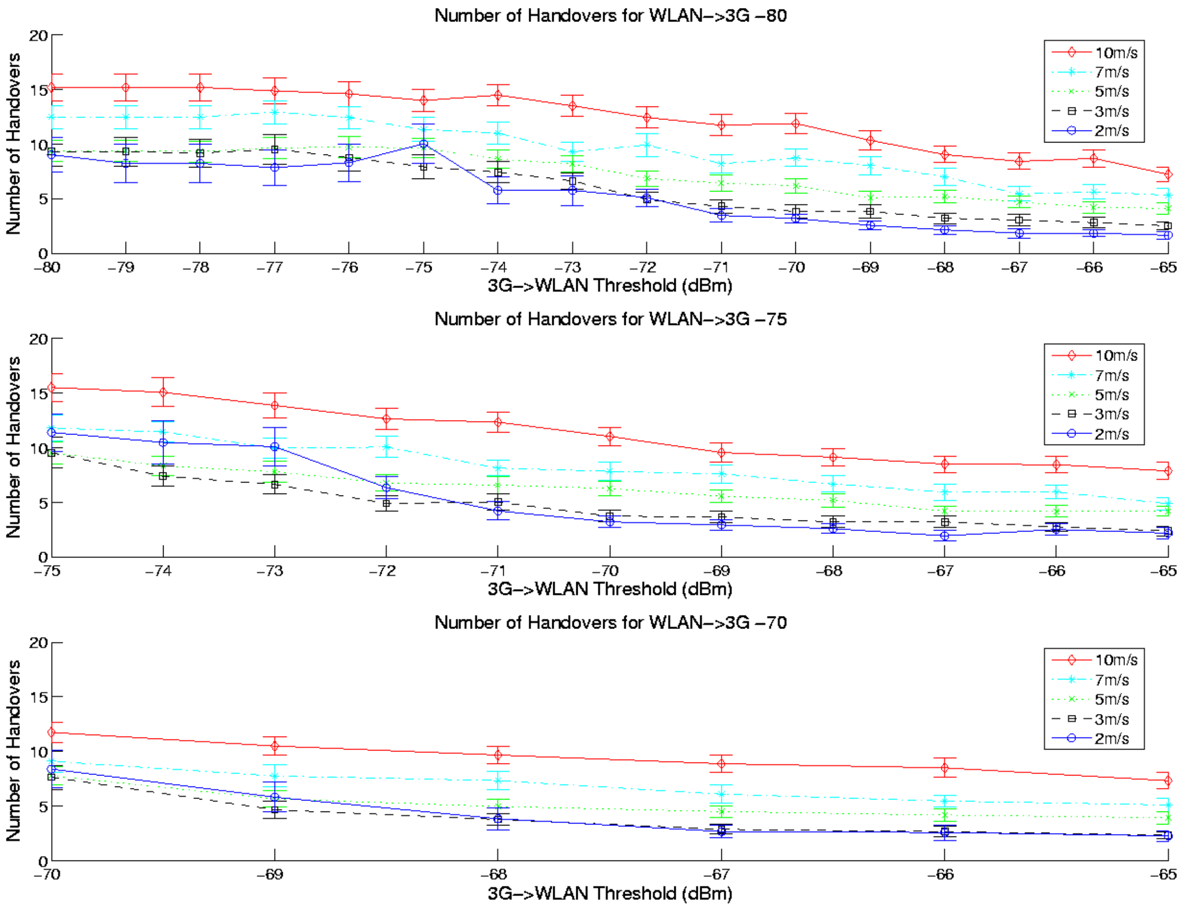

Figure 4.5: Number of Handovers for several speeds (RTT 3G 300ms)

Figure 4.6 shows the packet loss during handover. All the curves of figure 4.6 show a common behavior. Note that losses can be either due to signal variation or due to handover failure. As the $WLAN \rightarrow 3G$ threshold increases, losses (both due to signal variation and to a handover failure) are reduced. As a result this work defined the threshold configuration for zero packet loss, as the configuration of both thresholds in the Mobile Node with which a seamless handover is possible. The threshold configuration for zero packet loss varies for the different speeds. For speed value of 2m/s a configuration of $WLAN \rightarrow 3G = -70dBm$ and $3G \rightarrow WLAN = -70dBm$ is enough to provide zero packet lost. However, for the same threshold configuration and speed about 10m/s, on average 20 packets are lost. These values give insightful information for optimal terminal configuration and handover performance.

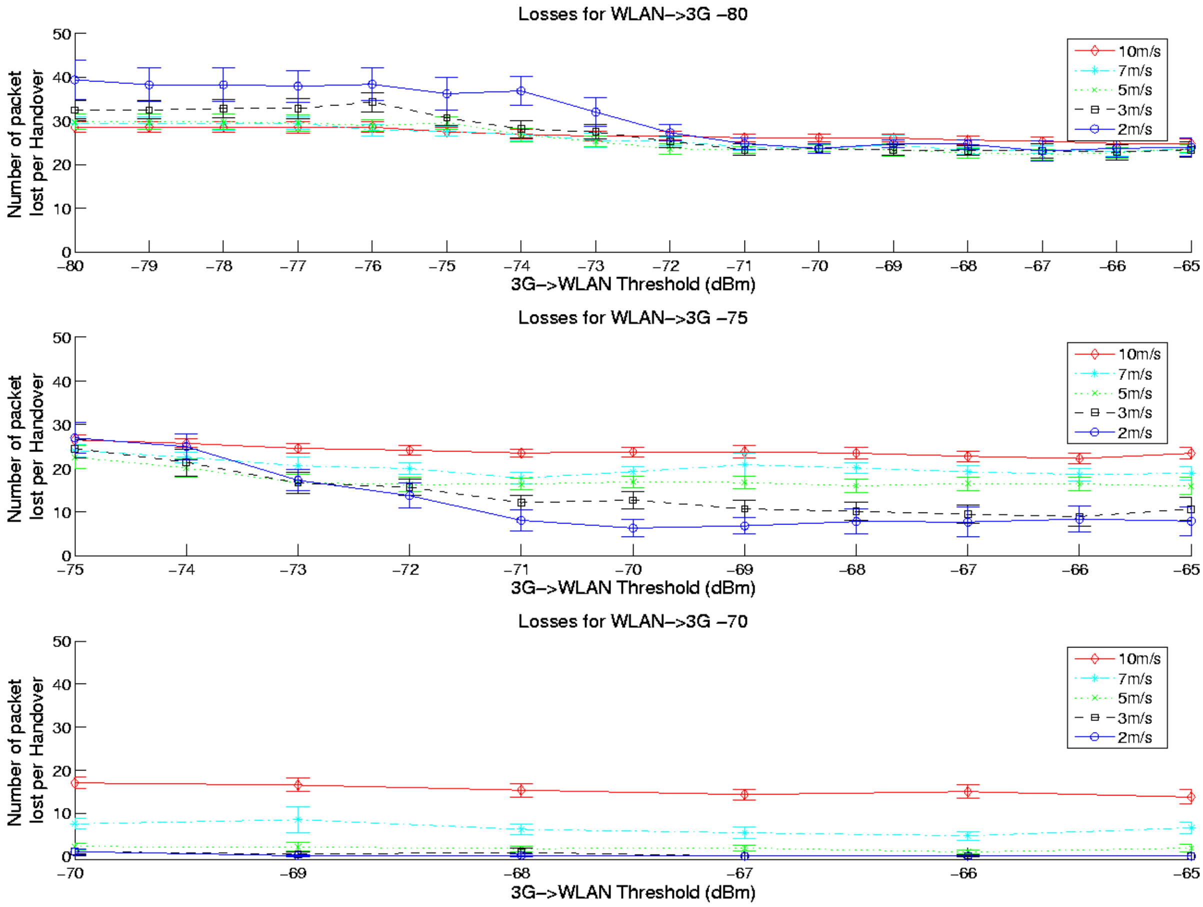

Figure 4.6: Number of Packet Lost for several speeds (RTT 3G 300ms)

4.3.2 Effect of the 3G channel RTT in the threshold configuration

To complete the study, an analysis on how the RTT of the 3G link affects the thresholds' configuration was done.

Figure 4.7 shows how the 3G channel RTT affects the Wireless utilization time, the number of handovers and the packet loss for a specific threshold configuration. Following typical values of RTT

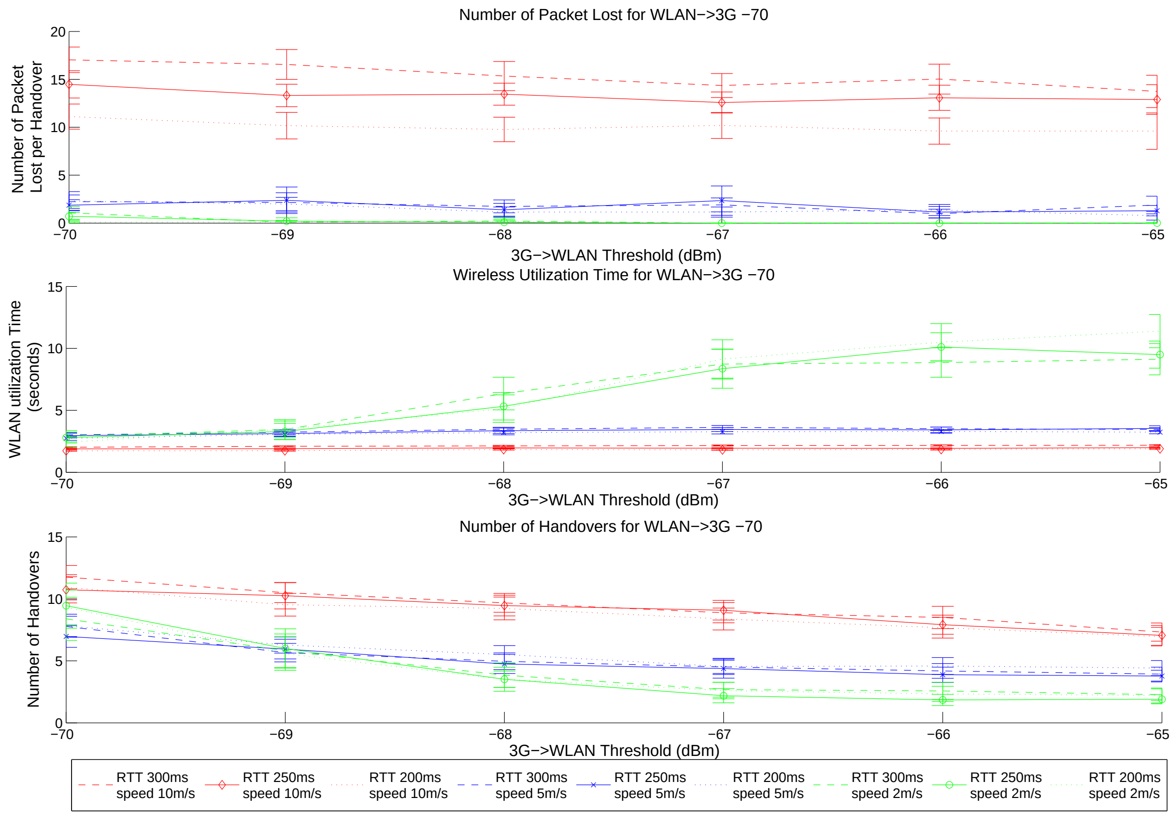

Figure 4.7: Wireless Utilization Time, Number of Handovers and Number of Packets Lost for several speeds and RTTs in the 3G Link

for an UMTS channel, which range between 190ms and 220ms[4], RTT was varied between 200ms and 300ms (i.e. the values used in the study are a worst case estimation).

Figure 4.7 shows that RTT affects neither the wireless utilization time nor the number of handovers performed. The major effect is in the number of lost packets. The reason is as follows. Since the RTT increases the time required to handoff to the 3G leg, the number of lost packets (due to WLAN signal level fading) increases accordingly. The effect is the same as if a less restrictive value for the $WLAN \rightarrow 3G$ threshold had been used.

4.3.3 Effect of the speed in the algorithm for measuring the signal level

As the handover algorithm is based on signal power thresholds and the signal level can typically vary a lot in indoor environments, the information reaching the MIH layer (e.g. RSSI each beacon interval) can be different in a relative short amount of time. Therefore, taking into account last samples, or a series of samples is not sufficient to derive the trend of the signal conditions. Thus, several approaches where proposed to infer the real trend of the signal (based on beaconing interval) against different speed conditions.

[4]Values measured with a commercial data card.

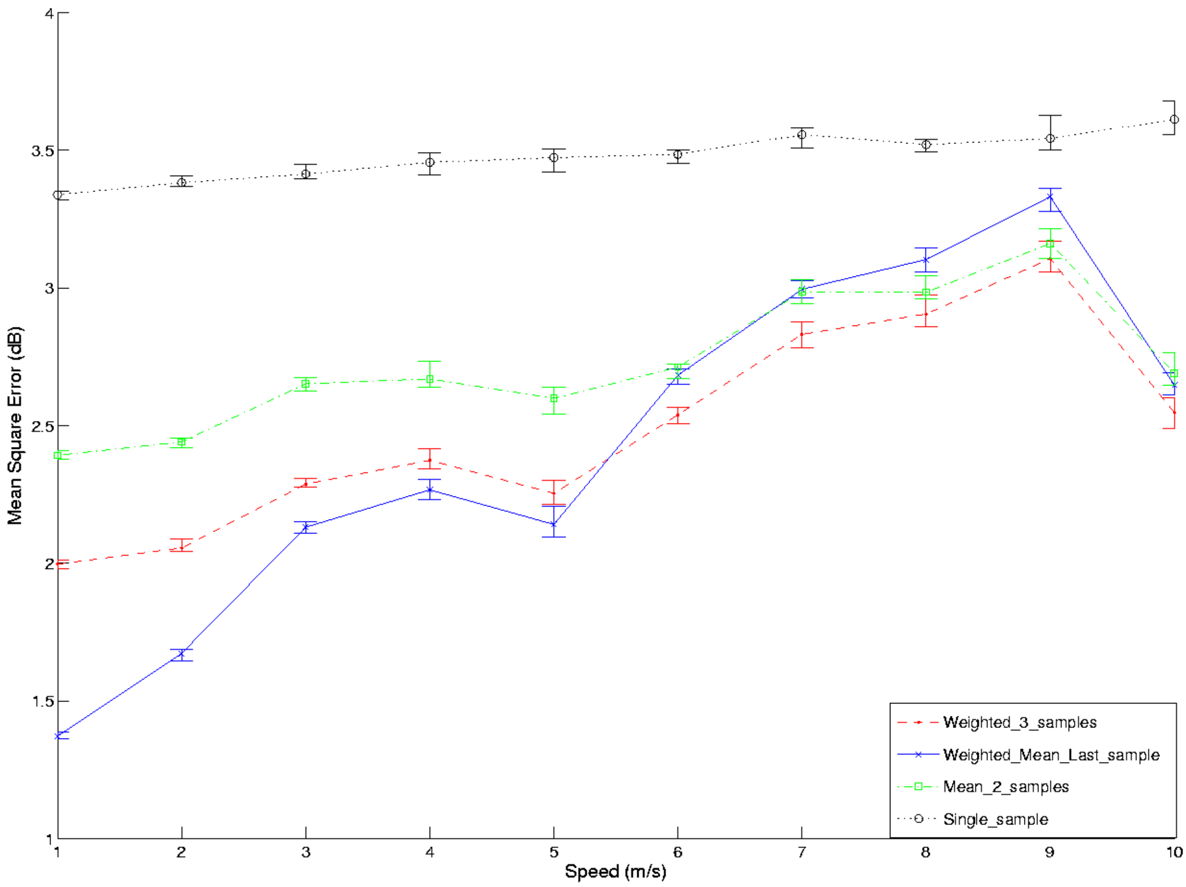

Figure 4.8: Mean Square Error of the signal behaviour prediction for different sampling algorithms

The different algorithms analyzed are:

- **Single Sample (SS):** The current value of the signal is the last beacon. $y[n] = x[n]$
- **Weighted mean (WM):** The current value of the signal is the value given by a weighted mean between the last beacon and the previous one. $y[n] = \alpha x[n-1] + \beta x[n]$
- **Weighted mean with the previous mean (WMPM):** The current value of the signal is the value given by the weighted mean between the last sample and the last mean. $y[n] = \alpha y[n-1] + \beta x[n]$
- **Weighted mean of three samples (WM3S):** The current value of the signal is the weighted mean between the last three samples.$y[n] = \alpha x[n-2] + \beta x[n-1] + \gamma x[n]$

Taking into account the four proposed algorithms, a simulation in Matlab was performed. For each algorithm, the optimal parameters of the weighted mean have been computed (trying all the combinations) taking into account several speeds. Figure 4.8 shows the Mean Error square obtained while evaluating the signal level for [SS], [WM],[WMPM] and [WM3S] techniques. From the results, it can be seen that for low speeds the algorithm *Weighted mean with the previous mean (WMPM)* outperforms the others, while for high speeds the algorithm *Weighted mean of three samples (WM3S)*

Speed	WMPM		WM3S		
	α	β	α	β	γ
1m/s	0.2	0.8			
2-3m/s	0.3	0.7			
4-5m/s	0.4	0.6			
6m/s			0.4	0.4	0.2
7-9m/s			0.5	0.3	0.2
10m/s			0.6	0.3	0.1

Table 4.1: Optimal parameters for the configuration of the *WMPM* and *WM3S* algorithms

gives the best performance. Based on these results, a combined approach dependent on terminal speed is recommended. Table I presents the optimal configuration for the two recommended algorithms.

As an additional example, the case of a terminal moving at 2 m/s could be considered. The optimal sampling technique is the WMPM and the optimal threshold configuration values are -75 dBm for $WLAN \rightarrow 3G$ and -70 dBm for $3G \rightarrow WLAN$. These values (not affected by the RTT) optimize Wireless LAN utilization time while providing acceptable packet loss rate. In a similar way, results can be derived from the graphs for other speeds and RTT.

4.4 Conclusions

The work presented so far identifies the terminal design to efficiently support MIHO in an heterogeneous environment. The simulative study proves that the IEEE 802.21 upcoming standard supports a common interface toward the access network providing a well defined set of commands and events and outperforms standard IP based mobility as identified in 4.1.2. This is a necessary step before illustrating, in the next chapter –the core of this work–, the solutions to support NIHO in an IP-based heterogeneous environment where the terminal architecture above presented is accordingly extended for MN-to-network protocol exchange.

Chapter 5

Implementing Simulations for IEEE 802.21 in Network Controlled Handovers

As referred to in section 2.4.2, NIHO provides improved resource allocation when heterogeneous wireless/wired access technologies are deployed.

To further show how NIHO technology can be efficiently supported in next generation mobile operator networks, modification and extensions to the ongoing IEEE 802.21 standardization effort have been proposed. Such extensions are based on the existence of a centralized entity and on the collection of relevant events from mobile devices. The proposed modifications are gathered in [54] explaining how the network configures the devices and how the intelligence triggers handover procedures.

The mobile node architecture recognizes the need for the abstraction layer design implemented by the MIH layer. Such layer is also implemented in the network side components to provide 802.21 protocol operation exchange. The Mobility Management Entity (MME) implements the intelligence required for centralized handover decision. The novelty of the approach consists in a scalable event-driven notification system the network can leverage for detecting critical conditions either in the terminal or in the network.

This section presents the core of the work focusing on network control for IP converged heterogeneous mobility, showing new simulation tests and results. Starting from the simulator implementation of 802.21 mechanisms at the terminal explored in the previous chapter, an extensive implementation effort was done to extend the simulation framework with extra code and functionality, taking into consideration terminal and network components for the support of NIHO. Implementation details can be found in the Appendix in sections B.1 and B.2.

5.1 Framework Design

This framework exploits the R3 IP based interface in IEEE 802.21, between the MN and the PoS (central entity), integrating the control signalling with Mobile IP signalling for data plane update. For simplicity (and due to its current industry relevance) this study is only applied across WLAN and cellular technologies.

In this scenario, global coverage from cellular technologies is always available, and enhanced coverage is available in multiple WLAN hotspots, a common situation currently. The mobile node typically performs a soft-handover (meaning that the new link is established before releasing the old one) between different interfaces, although this framework could be adapted to hard-handovers (in which the connection is set up through the new interface after closing the previous one in use). The network operational mode exploited here is Network Controlled/Initiated Handovers, assisted by the terminal (Mobile Assisted).

5.1.1 Mobile Assisted and Network Controlled/Initiated handovers

This operational mode places the handover decision mechanism in the PoS. When the MN reaches a WLAN cell and estimates there are favorable conditions, it will inform the network (PoS) of the new link detected, waiting for a confirmation from the network allowing or denying the execution of the handover procedure. Two separate sets of tests where executed under this operational mode. In the first test, the PoS assumes that resources at the target PoA are always available, not considering network load for the handover decision. This analysis will then assess the impact of the proposed IEEE 802.21 signalling compared to old scenarios where pure host driven mobility, which do not have the overhead of decision making signalling, is used. In a second stage a load balancing mechanism has been developed and tested, exploiting mobile node interface diversity for network optimization. The load balancing mechanism is explained in detail together with the signalling flow. The analysis of network controlled and initiated handovers will then show how network decisions can favourably impact terminal mobility, and which associated functionalities are required for these operations.

5.1.2 Signalling flows

Figure 5.1 presents the defined IEEE 802.21 signalling flow to perform a handover. The detailed list of parameters included in each message is presented in subsection 5.1.4.

5.1.2.1 3G⇒WLAN Handover

The signalling flow for the 3G⇒WLAN handover supposes a MN that is connected to 3G and is approaching a WLAN cell. As soon as an access point (AP) is detected as result of the Active Scanning procedure, the MIH Function at the MN receives a corresponding indication from the link layer and sends message (1) to the PoS, encoding the MAC address of the AP in a UDP packet. This message is followed by message (2), where information related to the change in signal strength is supplied to the PoS. The PoS is then able to verify information related to that target, such as the load value. Upon load evaluation (3) at the PoS, message (4) is received in the MN, which replies with message (5), informing if the handover is possible or not. Note that e.g. the handover target in the handover request might not correspond to the one the MN is located at, in case of network handover initiation (e.g. because of terminal mobility). The PoS, upon reception of this message, sends message (6). The MN processes this datagram in the MIHF, sending a local link command to the wireless interface, in step (7). Upon successful L2 association[1], message (8) is sent to the PoS. If the signal strength conditions are still favorable, the MN can execute a L3 handover (9) (a MIP registration) through the new link. Upon successful MIP registration, message (10) is sent to the PoS, which replies with message (11). Finally the MN is able to receive L3 traffic as result of the MIP binding procedure. Note that the difference between a soft and hard handover is only related with the moment when data is not further received through the old link, and does not affect the signalling flow.

[1] Please note that in the simulator an active scanning procedure is triggered prior to the attachment, in order to guarantee favorable radio conditions at the handover target.

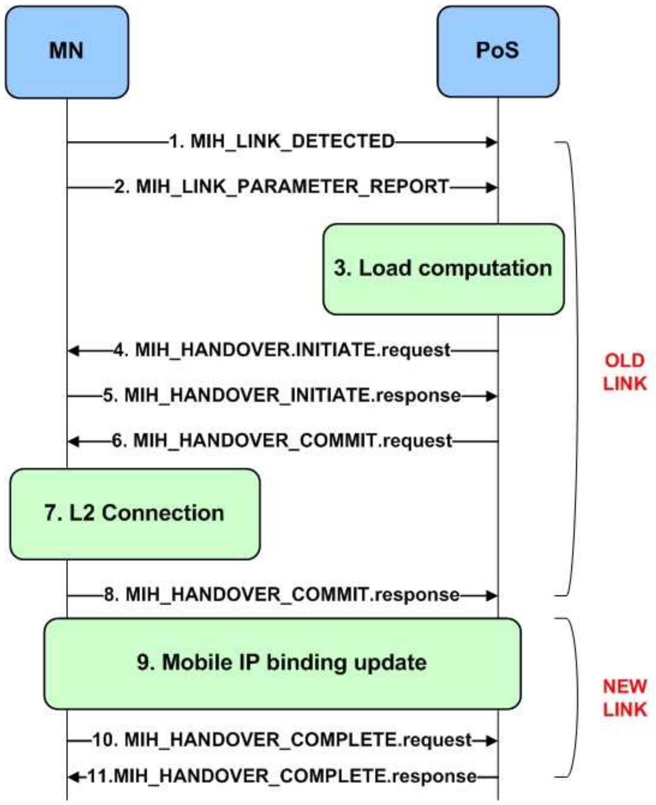

Figure 5.1: Handover Signaling for WLAN⇒3G and 3G⇒WLAN handovers

5.1.2.2 WLAN⇒3G Handover

This case supposes a MN associated to an AP, and the MIH Function continuously evaluating the signal level supplied by beacon messages. When the WLAN⇒3G threshold value is crossed, the MIH sends a Link_Parameters_Report (2) to the PoS, indicating deterioration of the received signal level. This will start a signalling exchange with the same messages and sequence as the 3G⇒WLAN handover, except for (1) MIH_Link_Detected that is omitted, since the 3G link is assumed always active (i.e. PDP context always active).

5.1.3 Load Balancing Mechanism

As stated before, a load mechanism has been implemented in the simulator. The use of this mechanism entails several changes in behaviour and signalling, presented in the following paragraphs.

Upon receiving indication from the MN of favourable link conditions, the PoS takes into account the load value of the handover target. Message 2 sent by the MN might not produce a reaction from the PoS, due to the target PoA being at high capacity. Thus a timer (to retransmit the Link Parameter Reports) is specified in order to refresh the PoS that the necessary handover conditions are still valid. The time value chosen for the timer is related to the RTT of the link, as recommended in the 802.21 specification.

For the load balancing procedure, each AP has an associated load value. The MN is also accounted in this load, affecting the value of the AP identified in the Link ID parameter of the respective MIH messages. An additional feature introduced by load balancing capabilities is the ability of triggering handovers for a MN when the load reaches the maximum value in a specific region of the WLAN network. This possibility can emulate the scenario of preferring the 3G coverage to a WLAN hotspot with a large load. In the considered scenario, high load in the AP means that video feeds could reach the MN with increased delay, packet loss, etc. So, when the MN is in WLAN and the load at that PoA is greater or equal than the maximum allowed value, the PoS sends an unsolicited handover initiate message to the MN, forcing a WLAN⇒3G handover.

Note that the reverse case is the usual behaviour of the handover process described in section 5.1.2. Through the use of events received from the MN, the PoS is aware of the MN being inside a WLAN cell. Hence, when the PoS verifies that the MN is connected to the 3G link and the load value of that AP presents itself good enough to admit a new entry (part of the operation in (3) of Figure 5.1), the PoS will initiate a 3G⇒WLAN handover, by sending message (4). Upon reception of this message, the MN will determine if the signal level is good enough for a handover.

In case a handover is both initiated by the MN and the PoS, to avoid concurrency problems, the event sent by the MN is ignored, and the handover initiated by the network continues normally.

5.1.4 Signalling Overhead

Given the reliance in the 802.21 signalling for the network operation, it is required to analyse the associated signalling overhead. IEEE 802.21 specifies a set of messages exchanged between the network and the terminal in order to perform a handover. The 802.21 frame is composed by a header and payload. The header consists of two parts: a fixed header which carries information related to the type of message and entity which is addressed to, and a variable header which helps in parsing the content of the payload. The first part is always present in any 802.21 message and has a fixed length of 8 bytes, while the second part carries information such as Transaction ID, Session ID or synchronization information and has a variable length.

This study supposes that the variable header is always present in the messages (worst case assumption) and its size is 8 bytes. The 802.21 message is completely defined in the payload, which is situated after the variable header. Inside the payload block, TLV encoding is used and the size of the payload block could be variable depending on the message and the parameters used. For each parameter, 5 more bytes should be added in order to complete the TLV format. Alignment to 32 bits is achieved by means of padding.

Table 5.1 specifies the messages and all parameters used in this study, with the respective sizes of each parameter. Although there is not any transport protocol defined yet for 802.21 datagrams, there are proposals that use UDP [41] (general design considerations are given in [40] based on a common set or requirements [42]). In this framework all the signalling has been performed over UDP/IPv6. For each packet a calculation of the packet size has been performed in the following way:

$$Length = IPv6 + UDP + FixedHeader + VariableHeader + TLV\ params \tag{5.1}$$

The signalling messages per handover sum 672 bytes, which, in the case of 3G to WLAN, 528 bytes correspond to signalling deployed through the 3G and 144 bytes correspond to signalling through the WLAN. In the case WLAN to 3G the numbers are reversed.

To get an understanding of the cost in terms of signalling when using 802.21, several calculations of the bandwidth used for signalling have been performed, taking into account the handover probability of our model. Studies like [55], argue that the average number of users in a 3G cell varies up to 52 users. For different numbers of users, the bandwidth used for signaling can be calculated and is depicted in table 5.2, considering simulation runs of 2000 seconds.

MIHF Protocol Message	Parameter Name	Type	Size
MIH_LINK_DETECTED	Link ID	Network type	4
	MacNewPoA	MAC Address	6
MIH_LINK_PARAMETER_REPORT	LinkParameterType	Link Quality Parameter Type	1
MIH_HANDOVER_INITIATE.request	Handover Mode	Handover Mode	1
	SuggestedMacNewPoA ID	Mac Address	6
	CurrentLinkAction	Link Action	4
	SuggestedNewLink ID	Network Identifier	4
MIH_HANDOVER_INITIATE.response	Handover ACK	Handover Mode	1
	Preferred Link ID	Network Identifier	4
MIH_HANDOVER_COMMIT.request	NewLink ID	Network Identifier	4
	NewPoAMAC	Mac Address	6
	CurrentLinkAction	Link Action	4
MIH_HANDOVER_COMMIT.response	OldLinkAction	Link Action	4
MIH_HANDOVER_COMPLETE.request	Handover Status	Status	1
MIH_HANDOVER_COMPLETE.response	ResourceStatus	Resource Retention	1

Table 5.1: Messages and associated parameters (size in Bytes).

	2m/s		5m/s		10m/s	
N° User	WLAN	3G	WLAN	3G	WLAN	3G
20	6.6±0.6	24.4±2.2	27.7±1	101±3	40.9±2	150±7.6
40	13.3±1.2	48.8±4.5	55.3±1.9	203±7	81.9±4.2	300±15

Table 5.2: Signalling Bandwidth cost in Bytes/sec in function of mobile node speed in m/sec

In this table, it can be seen that the signalling load increases with the number of users and their speed of movement, but in all cases, signalling load remains very low. In the worst case (40 users and 10 m/s) the required signalling corresponds to 300 bytes/second in average, delivered through the 3G link; and 82 bytes/second, delivered through the WLAN. This result corresponds to handovers from 3G to WLAN. The inverse case (WLAN to 3G) has similarly corresponding values.

It can be pointed out that the signalling specified in IEEE 802.21 is loading the network very lightly and is enough to support a high number of users performing handovers between different technologies like WLAN and 3G. This supports the intention of exploiting 802.21 MIH functionalities to aid heterogeneity mobility.

5.2 Simulation Modeling

The study was conducted by simulating the movement of a MN attached to a 3G network and performing several handovers between 3G and WLAN hotspots, varying terminal speed and coverage threshold values.

The simulation scenario considers wide space with indoor characteristics (such as an airport) in which the user can move at different speeds and it closely follows the network scenario mentioned in section 5.1. It consists of an environment with a partial area of non-overlapping WLAN cells[2] and full coverage of 3G technology. The WLAN coverage is supplied by Access Points, each connected to an Access Router. The scenario also features a Home Agent for the MIP Registration process,

[2] The setup features four access points distributed in a square area of 500X500 meters.

an audio server which streams audio traffic to the MN[3], and the PoS which is the central network entity that exchanges MIH messages with the MN. This adds the network part of the IEEE 802.21, under standardization, to the model studied in the previous chapter, thus creating a framework suited to model Network Initiated and Assisted handovers. Through the rest of this section several details of the model and the specification of the algorithm which conform the PoS and MN behavior, are provided.

This simulation scenario is similar to the one presented in [56] and [57] with the difference that in those contributions only Mobile Initiated Handovers, and without any network control, were considered. As a consequence there was neither the concept of central entity (the PoS) controlling mobility, nor IEEE 802.21 signalling over the air between the mobile node and the network.

Movement Pattern

The movement pattern selected is the Random Waypoint Mode. The MN moves between uniformly distributed waypoints, at speeds of 2m/s, 5m/s and 10m/s targeting to model speed scenarios that will be the usual worst case in WLAN environments, including the border between WLAN and 3G (the focus of our simulations). In section 5.4, the effect of higher speeds is also studied.

WLAN Model

The WLAN Model used is the one implemented in OMNeT++ based on free space losses with shadowing and a variable exponential coefficient. Each simulation was run with 3G⇒WLAN and WLAN⇒3G thresholds varying between -75dBm and -65dBm.

Load Factor

For the load balancing optimization, a birth-and-death Poisson process is used, considering a maximum number of clients per AP. Different user inter-arrival rates where simulated varying network load from 50% up to 100% of the maximum system capacity.

The 3G Channel Model

The 3G channel has been modeled as a PPP channel with a connection time of 3.5 seconds, disconnection time of 100 ms, bandwidth of 384 kbps (downlink) and variable delay of 100 to 150 ms per way[4]. Although the above model takes into account the connection time, in the executed simulations it was assumed that the PDP context is always active, so the value of the connection time does not have any impact. As per the work analyzed in the previous chapter, these simulations are based on the following two assumptions i) full 3G coverage and ii) 3G link always on,which are realistic assumptions in typical scenarios.

Metrics used in the study

The main focus of the simulation work was to verify that the introduction, in a threshold based handover algorithm, of the IEEE 802.21 signaling that enables network control, does not hinder the ability to achieve a good use of the wireless cells. For exploring this issue the following parameters where used:

[3]The traffic studied is a downstream audio, with a packet size of 160 bytes at application layer and inter-arrival packet time of 20 ms (83 kbps). Notice that usual VoIP codecs generate bit rates around 80 kbps and therefore their traffic pattern is very similar to the simulated one.

[4]Measurements have been taken with a commercial 3G data card.

- Mean percentage of L2 handover without MIP registration (failed handovers)
- Mean number of 3G⇒WLAN handovers
- Mean number of WLAN⇒3G handovers
- Mean wireless utilization time

Regarding the first metric, a failed handover is a situation in which the mobile node detects the WLAN cell and starts the signalling procedure in figure 5.1 but, after receiving message (6) the signal level never goes over the 3G⇒WLAN threshold, and the procedure is not completed, in particular a layer three registration to send the traffic to the WLAN interface does not take place. Notice that this situation does not imply any connectivity problem, as communication continues normally using the other interface. The second and third metrics are related to the mean number of 3G⇒WLAN and WLAN⇒3G handovers, respectively. Lastly, the mean wireless utilization time is also accounted.

Extended Terminal Architecture for NIHO support

The terminal's architecture includes a subset of the Media Independent Handover Protocol defined in [1]. This work focuses on the impact of the required signalling to perform handovers while mobile terminals move at different speeds, thus MIH capability discovery and remote registration are supposed to already have occurred.

The handover algorithm in [56] reacts to events resulting from the analysis of the signal strength in the WLAN interface. A MIH implemented in the MN supplies triggers to a local decision engine, based on 3G⇒WLAN and WLAN⇒3G thresholds, possibly resulting in a handover. This chapter describes the implemented complement to this algorithm with MIH signalling between the terminal and the PoS. Figure 5.2 depicts the message exchange intelligence implemented in the simulator, in the MIH layer at the MN. This message exchange allows the MN to supply fresh information about current link conditions to the PoS, as well as to receive remote commands for handover initiation. The message exchange is triggered upon signal level threshold crossing and generates local link events. These events are 1) LINK_DETECTED when the terminal detects a new WLAN cell, 2) LINK_-PARAMETERS_CHANGE when the received signal level crosses a configured threshold, and 3) LINK_UP that indicates a successful L2 connection establishment. These events are collected in the MIHF of the MN and conveyed to the MIHF in the PoS.

The first two events supply to the PoS an indication that favorable handover conditions are available to the MN, and may result in signalling between the two entities for a handover initiation. When the necessary message for handover initiation is received from the PoS, the MN is able to perform the L2 handover. The terminal keeps analyzing the signal level and when a configured 3G⇒WLAN threshold is crossed, a layer three handover can occur. In this phase, the MIP signalling takes place updating in the HA the new MN's CoA. Due to the configured 3G⇒WLAN threshold, and also to the movement of the node and the delay caused by the signalling, a layer two handover might not lead to a Mobile IP registration (this is one of the metrics of the simulation model, which is extensively studied in section 5.3). Since the analysis focus is on inter-technology make-before-break handovers, the MN will attempt to establish the new link before releasing the old one. When the MN is connected to the WLAN, and the MIH Function verifies that the received signal strength is not favorable anymore, a WLAN⇒3G is triggered. Thus, the MN starts the MIH signalling to the PoS, potentially initiating a handover to the 3G link.

While evaluating the more suitable algorithm for the MN, it was decided to perform the MIH signalling once the MN reached the WLAN cell. Thus, when the signal level crosses the 3G⇒WLAN

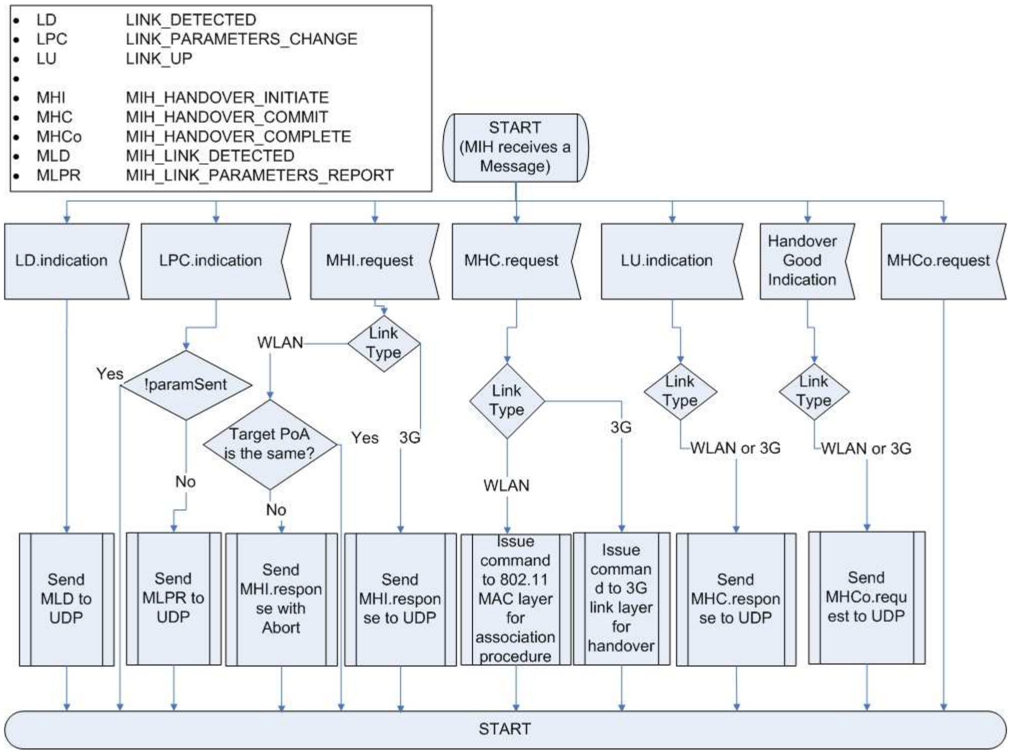

Figure 5.2: MIH Intelligence at the MN

threshold, MIP signalling is sent to complete the layer 3 handover. The use of this model leads to higher MIH signalling load upon cell detection, but avoids possible delay for signalling completion between layer two link detection and the layer three handover processes. A study comparing two operational modes regarding 802.21 signalling opportunities can be found in section 5.5.

PoS Design

The PoS is a network entity whose MIHF is registered to the MN's own MIHF, receiving subscribed events. Through the received messages, the PoS tracks down the terminal's position and the quality of its received signal strength. Then, the PoS can supply a remote command for handover initiation depending on the load value in that AP. The implemented PoS intelligence depicted in figure 5.3. This is implemented as a network node with a full 802.21 MIHF stack, having the ability to send and receive MIH signalling encapsulated in UDP packets [19], and a decision engine for handover execution.

The PoS also has two operational modes depending on the active simulation scenario, where load processing can be active or not. In this last case it always supplies an affirmative handover command when called.

5.3 Evaluation

This section first presents the scenario where no admission control mechanism is applied. Figure 5.4 depicts the percentage of failed handovers. Three speeds have been considered namely, 2, 5 and 10 m/s targeting indoor scenarios. From the graph it is visible that by varying the threshold 3G⇒WLAN

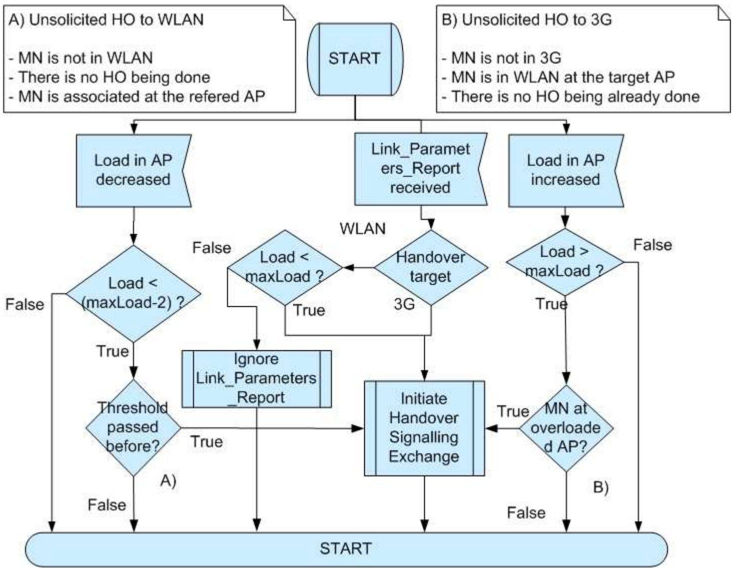

Figure 5.3: PoS Intelligence

from -75 up to -65 dBm the percentage of failed handovers (as defined above) increases to almost 65% in case of 10 m/s. The curves follow a similar shape for 2 and 5 m/s. As can be noted, the curves show a trend to increase while the 3G⇒WLAN threshold value is increased.

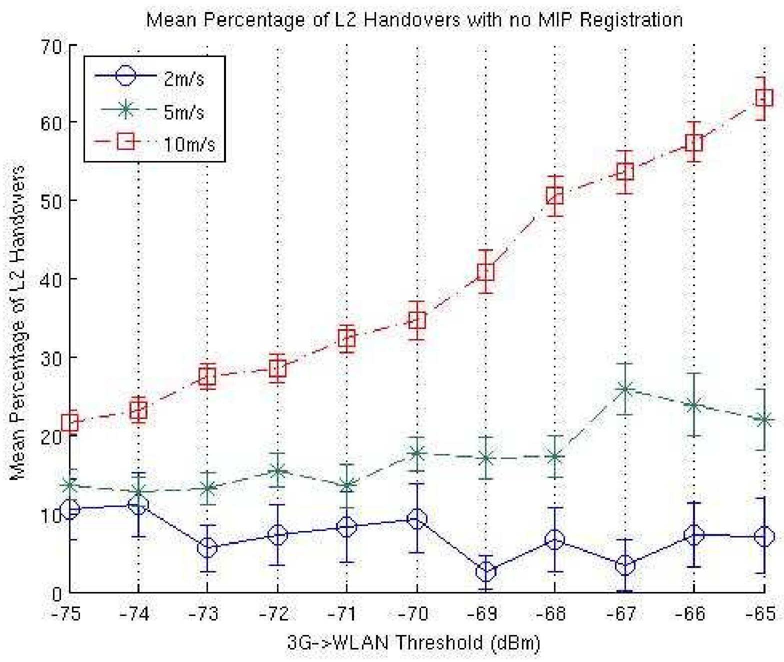

Figure 5.4: Mean percentage of layer two associations not followed by a layer three handover when WLAN⇒3G threshold configured at -75 dBm

When the mobile node detects a WLAN cell, it starts the signalling procedure of figure 5.1. After receiving message 6, the mobile node checks the signal level received from the WLAN AP and continuously evaluates the signal until the level crosses the 3G⇒WLAN threshold for continuing with the signalling. If the signal level never reaches a value over the 3G⇒WLAN threshold, we have a failed handover. This can happen naturally because of the mobility pattern. The mobile approaches the WLAN cell, but because its movement direction or a sudden movement change, it never reaches the position in the cell where the signal level is above the threshold. Of course, as the 3G⇒WLAN threshold is higher, this happens more often, as can be observed in figure 5.4. Faster speeds also increase the number of failed handovers, because in more occasions the mobile is not enough time in the zone inside the threshold.

An important point is the impact of the delay introduced by the required signalling in this procedure. Without the signalling to enable network control (figure 5.1), the mobile node is ready to perform the handover immediately after detecting the WLAN cell. With the signalling, we introduce a delay (the time between message 2 in figure 5.1 and receiving message 6) in which, even if the signal level crosses the threshold, the mobile node cannot perform the handover because it has to wait to complete the signalling with the network. If the delay introduced by the signalling is larger than the time needed to cross the 3G⇒WLAN threshold, the handover is delayed or in the worst case it even doesn't occur. This issue is explored in table 5.3 in which the delay from sending message 2 to receiving message 6, and from sending message 2 to finishing step 7, is compared for different speeds and 3G⇒WLAN thresholds. The signalling delay is much lower than the time needed to cross the threshold and completing step 7, showing that the signalling does not interfere with the handover performance. With this, it is possible to argue that the mobile node to network communication is suitable both from a signalling overhead point of view (table 5.2) and from handover performance point of view (table 5.3).

Threshold / Speed	-75dBm	-72dBm	-69dBm	-66dBm	-65dBm
Time from sending message 2 to receiving message 6 (3G⇒WLAN)					
2m/s	0.43±0.0002	0.43±0.0002	0.43±0.0002	0.43±0.0005	0.43±0.0002
5m/s	$0.422\pm4.5\text{x}10^{-5}$	$0.422\pm4.8\text{x}10^{-5}$	$0.422\pm9.8\text{x}10^{-5}$	$0.422\pm5.5\text{x}10^{-5}$	$0.422\pm4.1\text{x}10^{-5}$
10m/s	$0.421\pm2.8\text{x}10^{-5}$	$0.421\pm2.8\text{x}10^{-5}$	$0.421\pm3.03\text{x}10^{-5}$	$0.421\pm3.4\text{x}10^{-5}$	$0.421\pm3.3\text{x}10^{-5}$
Time from sending message 2 to finishing step 7 3G⇒WLAN)					
2m/s	13.6±0.4	20.6±0.8	25.5±1.3	27.1±1.5	28.9±2.2
5m/s	4.4±0.07	6.1±0.1	7.6±0.2	8.5±0.2	9.0±0.3
10m/s	2.1±0.03	2.9±0.05	3.7±0.07	$4.1\pm0.1\text{x}10^{-5}$	4.3±0.08

Table 5.3: Time required in performing signaling depicted in figure 5.1 for selected 3G⇒WLAN thresholds.

Figure 5.5 depicts the mean number of layer three handovers obtained by varying the 3G⇒WLAN threshold. The impact of the speed affects the metric in different ways depending on the considered configuration. At the value -75 dBm the number of handovers is quite large especially considering a high mobility level, while it decreases and converges for greater values of the threshold. The decay in the slope of the different speeds is related with the failures of performing the layer three handover shown in figure 5.4. The graph shows how the values tend to converge, when the 3G⇒WLAN threshold is increased. The graph presenting the number of handovers from WLAN to 3G is symmetric due to the scenario symmetry. It is interesting to note that the closer the mobile node to the access point, the lower the chance of having complete handovers. This is complementary to the previous graph, as the metric is mostly affected by the mobility pattern and not from the signalling required for mobile to network communication.

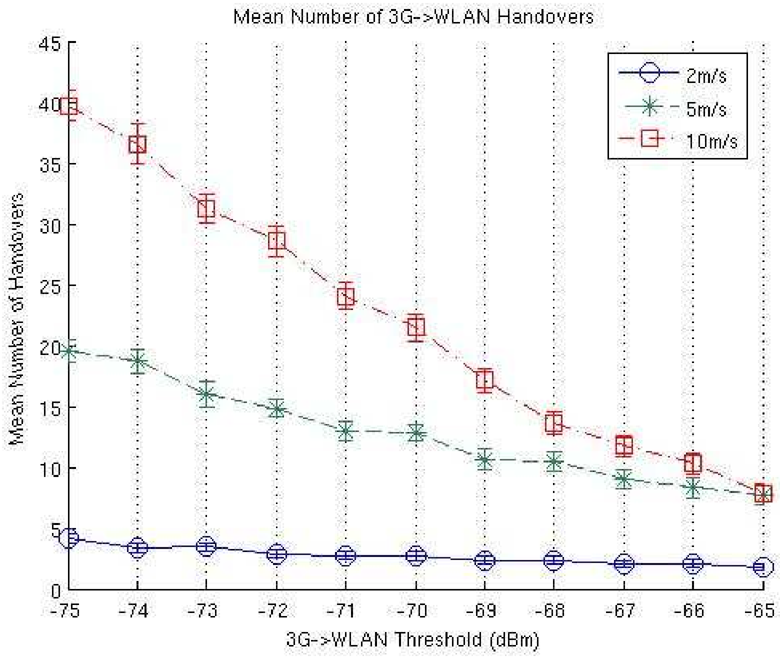

Figure 5.5: Mean number of 3G⇒WLAN handovers when the WLAN⇒3G threshold is configured at -75dBm

Figure 5.6 shows the mean wireless utilization time according to the three different speeds. The general observed behaviour is a flat response with the increase of the 3G⇒WLAN threshold. As the

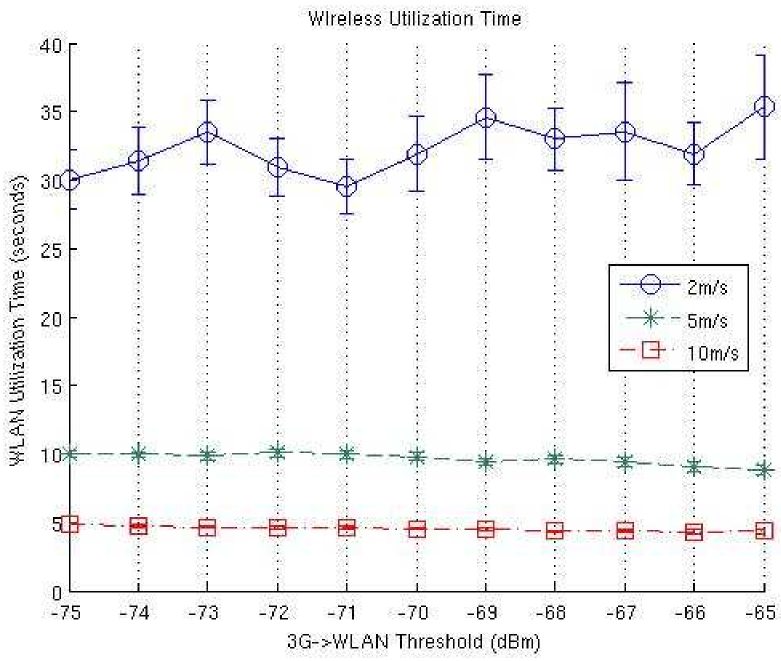

Figure 5.6: Mean wireless utilization time (units of time per handover)

primary goal of this study is the maximization of the wireless utilization time, and thus to reduce the number of handovers which do not result in a long term stay inside the cell, figure 5.6 demonstrates that the signalling does not impact the mean wireless utilization metric. In fact, the relative magnitude between the different lines shows that the metric is mostly impacted by the time the user resides in the wireless cell, which result in a higher utilization time at lower terminal speed. This conclusion further supports the explanation of figure 5.4 where the mobility pattern represents the dominant effect on the system.

The results above presented demonstrated that if values in table 5.3 are verified, the cost of mobile to network signalling for network controlled and initiated handovers is negligible. This is an insightful result, especially considering environments (e.g. WLAN hotspots) where network controlled mobility is not yet considered as core technology to improve both user experience and network resource usage. Next follow the the results obtained for the load balancing scenario defined in 5.1.3 taking as references figures 5.4, 5.5 and 5.6.

Figure 5.7 represents the number of failed handovers as defined above, while load balancing is applied. The behavior is similar to the one in figure 5.4, since the framework for network initiation

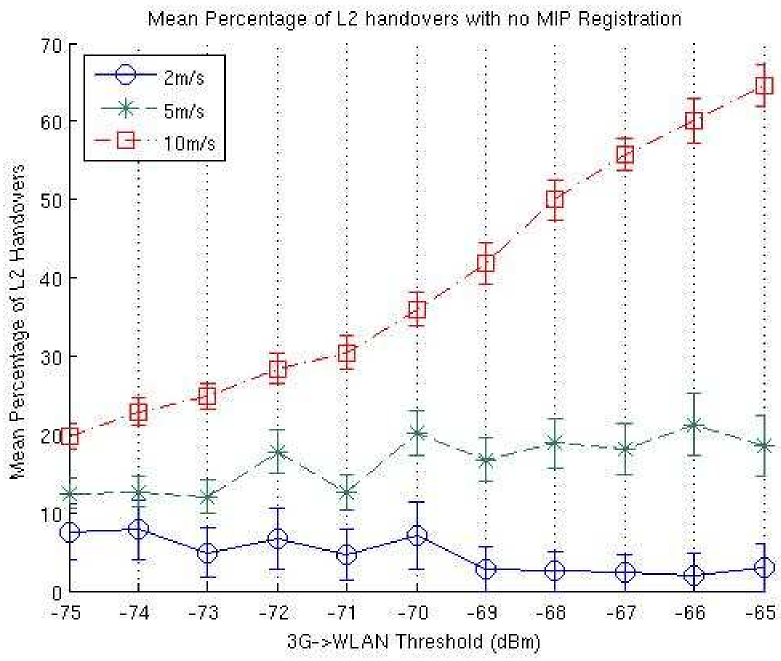

Figure 5.7: Mean percentage of layer two associations not followed by a layer three handover when WLAN⇒3G threshold is configured at -75 dBm. Load balancing scenario.

accounts the terminal for the most up to date report information. The percentage of failed handovers due to wrong location report is around 3%, which seems an acceptable result. Figure 5.8 accounts for the number of handovers to the WLAN. The metric is directly impacted by the admission control mechanism and the load generated on the different access points, where a slightly smaller number of handovers can be verified between figure 5.8 and figure 5.5. It is worth noticing how the load balancing mechanism is not affecting as much lower speeds as 2m/s and 5 m/s as it is affecting 10 m/s. The values for these two lower speeds are not changing in a noticeable way between figure 5.8 and figure 5.5. In this perspective, the results prove the validity of the approach making load balancing

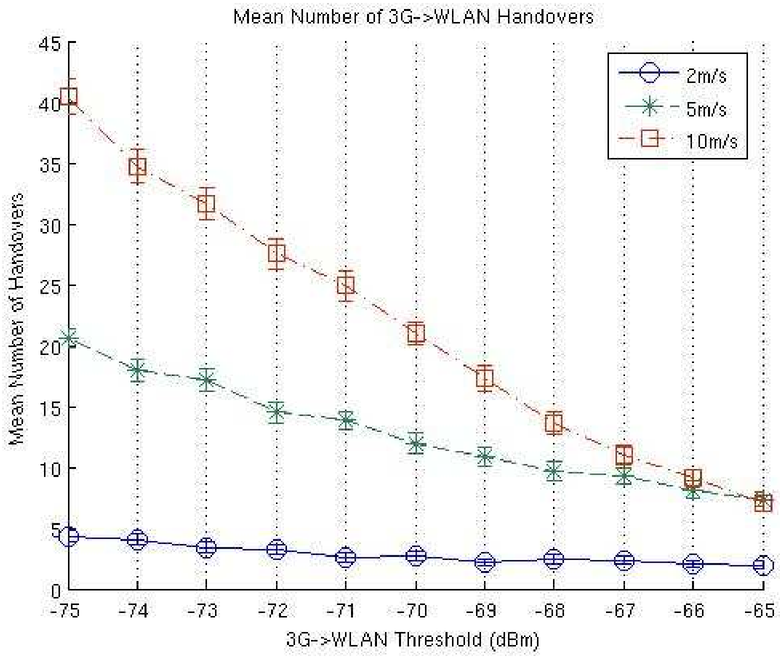

Figure 5.8: Mean number of 3G⇒WLAN handovers when the WLAN⇒3G threshold is configured at -75 dBm. Load balancing scenario.

scenarios attractive from an operator point of view.

Table 5.4 compares the wireless utilization time with and without load balancing, considering capacity usages of 50% and 100%. By comparing these results, it would be expected that the wireless utilization time decreased, but as can be noted, the utilization time is not decreasing equally for all speeds, and the 10 m/s speed is the most affected one. This behaviour can be explained with the fact that the help of network initiated handovers reduces the overall number of performed handovers and at the same time increases the overall wireless utilization time. This is a desirable feature in next generation networks where minimizing the network overhead is a must, especially in last hop wireless channels.

Finally and for completeness, evaluations of the impact of RTT were done to verify its effect on the 3G link. Simulations where RTT values varied between 200ms and 300ms showed only quantitative differences, maintaining the general behaviour of the previous graphs.

Speed (m/s)	No Load Balancing	Load Balancing 50% capacity	Load Balancing 100% capacity
2	32,4s	30,9s	25,9s
5	9,65s	9,46s	9,05s
10	4,53s	4,55s	4,45s

Table 5.4: Wireless usage with and without load balancing

5.4 Impact on 4G Design Architectures

The results presented in the previous section validate the framework design showing the feasibility of a new approach for mobility and handover management. Specifically the IEEE 802.21 signalling, while introducing minimized network overhead, leads to optimal network control of terminal mobility. The comparison of simulation results with and without network load knowledge shows a negligible impact on the chosen metrics. However, when considering future 4G networks and wide scale deployments there are some further issues that should be accounted. That is, the configuration of optimal thresholds for WLAN⇒3G handovers is critical to avoid signalling packet loss and should be complemented with accurate methods for out of cell detection. These issues are briefly described in the following sections.

5.4.1 Optimal configuration for WLAN⇒3G Handover

The case analyzed is the worst case condition when the terminal performs handover from the wireless LAN to the 3G link. Since the 802.21 signalling is always performed through the current link there might be conditions in which the signalling could not be completed, and added mechanisms are required as fall back solutions. Although a transport protocol will introduce ACKs and retransmission of the lost packets, the effects shown in this section must be taken into account or the transport reliability will introduce undesired delays.

Figure 5.9 shows the effect of the WLAN⇒3G threshold on the signalling between the MN and the PoS. The picture shows, for each simulated speed, the number of signalling failures to perform handover from the WLAN link to the 3G link. The results indicate that at high speeds (10m/s) we obtain a high mean number of interrupted/failed signalling flows with the PoS.

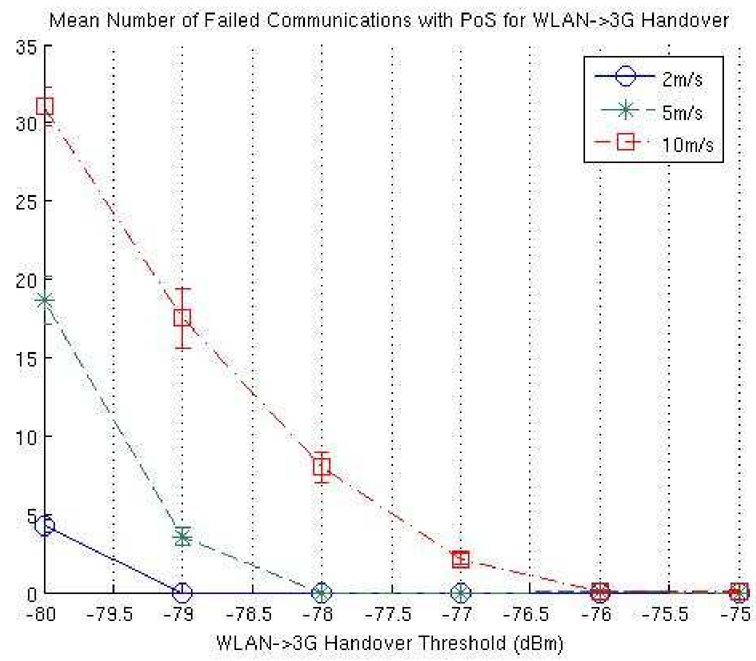

Figure 5.9: Effect of the -80 dBm threshold on handover signalling

This number increases while decreasing the WLAN⇒3G threshold. This behaviour can be explained as the result of the MN going out of the cell before the signalling flow ends. As the WLAN⇒3G threshold increases (in dBm) the signalling between the PoS and the MN starts before and the probability of going out of the cell decreases. Regarding the MIH functioning on interrupted signalling, this occurrence falls back on transport issues, which incorporate delay and loss of messages (as stated in [56]).

MIH Functions existing at the MN and PoS can optionally implement the optional Acknowledgement mechanism. In the case of interrupted signalling, this event would be dealt as if messages where lost. Also, the behaviour from the terminal in case a LINK_DOWN is received in the MIH is implementation dependent. For example, upon connection to a new available link, the MIH at the terminal can send a MIH message to the PoS requesting a handover rollback for freeing resources previously reserved for the handover that failed. This behaviour can free the resources faster than waiting, for example, for a timeout.

5.4.2 Out of Cell Mechanism Detection

The load balancing mechanism studied previously is based on the assumption the PoS is aware of the current L2 location of the terminal. We propose to exploit 802.21 capabilities to update the PoS with the information on the current location. The mechanism bases on the fact that the terminal via internal state machine can determine with the help of the MIH function whether he is approaching a WLAN cell or he is leaving a cell previously visited. Through the use of 802.21 events, more concretely the LINK_DETECTED event, it is possible for the MIHF in the terminal to gain knowledge of which WLAN cells it is able to detect (through AP MAC address contained in 802.11 Probe messages). Since the terminal can determine with acceptable accuracy the RSSI from the visited cell, we propose to convey this information to the PoS to enable better target choice while performing load balancing. The rational behind is as follows. In order to successfully move terminals form one cell to another to optimize network load the network has to determine the current location of the terminal. Indeed, the selected cell should also be visible from the terminal point of view. Nevertheless the freshness of that information is crucial in the decision process although a trade off between freshness of the information and signalling overhead in the network must be considered.

5.4.3 Signalling Breakdown in Very Fast Node Handovers

The approach described in this paper is based on the assumption that the IP layer is the common convergence layer across heterogeneous technologies. In case the signalling is applied to devices integrating broadband wireless access technologies such as WLAN and WiMax, it would be desirable to identify what are the upper bounds in terms of stability and reliability not affecting performance of the handover procedures. A specific scenario was analyzed, featuring one single WLAN cell that the mobile node crosses following a straight line. This movement pattern is similar to automotive/train scenarios where vehicles/trains can move only along predefined paths. The experiments have been performed for selected thresholds and letting the mobile node moving with increasing speeds, up to 35 m/s. This setup is deemed sufficient for investigating how the threshold based algorithm and 802.21 signalling perform in such speedy scenarios.

The graph in figure 5.10 presents the result of the study. In this graph the highest speed at which handovers finish successfully for different 3G⇒WLAN thresholds is represented. As can bee seen, it shows that the performance of the system rapidly decreases crossing the -65 dBm threshold. This is the expected behavior, as the failures are function of the speed. It should also be noted that the study in figure 5.10 considers the results shown in figure 5.9 where the optimal threshold configuration guaranteeing no packet loss due to WLAN signal fading is configured at -75dBm. This study completes

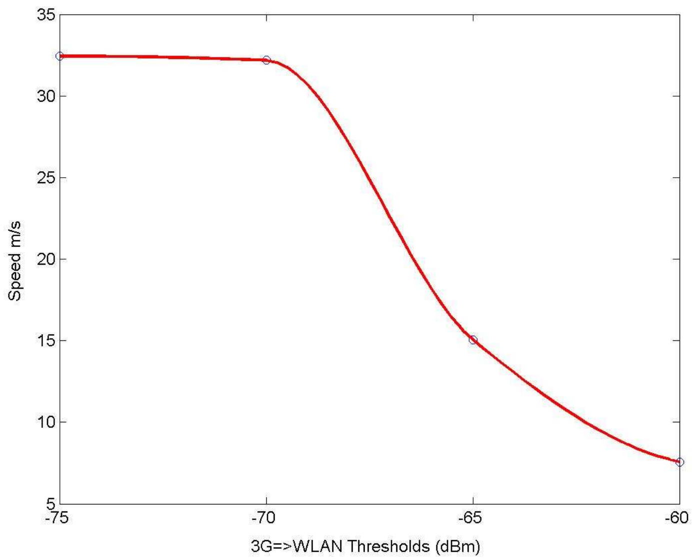

Figure 5.10: Interpolation of values showing system breakdown based on the speed.

the results presented in the previous section giving insights on the applicability of the technology in speedy scenarios providing wireless broadband access.

5.5 Impact on Terminal Design

Figure 5.1 describes the signaling exchange between the network and the terminal, for the target case of 3G and WLAN technologies. In the following section, based on [2], a focus was placed on the terminal design (including L2 functionalities) and its integration with the L3 signaling.

5.5.1 Thresholds

The terminal's intelligence relies on several thresholds across the signal strength evaluation, as can be seen in figure 5.11. While connected to the 3G link, the terminal is able to collect probe responses and beacons from the access points, evaluating the received signal strength indication (RSSI). Two thresholds are defined, namely the association threshold and the 3G⇒WLAN threshold. The first one refers to the mean signal strength required for the terminal's intelligence to decide to connect to an access point. The second one refers to the mean signal strength required for the terminal to decide that a successful 3G⇒WLAN handover is possible. Furthermore, when connected to the network through the WLAN link, a WLAN⇒3G handover threshold is defined to determine when the signal strength conditions require a handover to the 3G link. It is worth noting in figure 5.11 that the 3G⇒WLAN threshold is defined as being greater (in dBm) than the WLAN⇒3G threshold for zero-packet loss, as analyzed in the configuration of [56] and [57]. Also, the association threshold is defined as being lower (in dBm) than the WLAN⇒3G threshold.

5.5.2 Operational Modes

Depending on when (time-wise) the L3 signaling is triggered, two different operational modes have been implemented in the simulator, hereinafter referred as Operational Mode A (OM.A) and Operational Mode B (OM.B). These two modes execute the signaling model presented in section III, but differ in the timing of occurrence of certain messages, more specifically in the point at which the IEEE 802.21 remote signaling related to L2 connection execution is started. In OM.A the signaling to the PoS is triggered at the configured 3G⇒WLAN threshold, and in OM.B the signaling to the PoS is started at a fixed value of -80dBm, the association threshold. Both Operational Modes have been implemented to compare the effect of how the timing of the signaling messages affects handover procedures. Figure 5.11 visualizes the correlation between events (WLAN cell connection) and signaling triggers to the network. It is important to note the splitting of the signaling between cell detection, and associated RSSI report, and the signaling for effective L3 handover, including Mobile IP binding update.

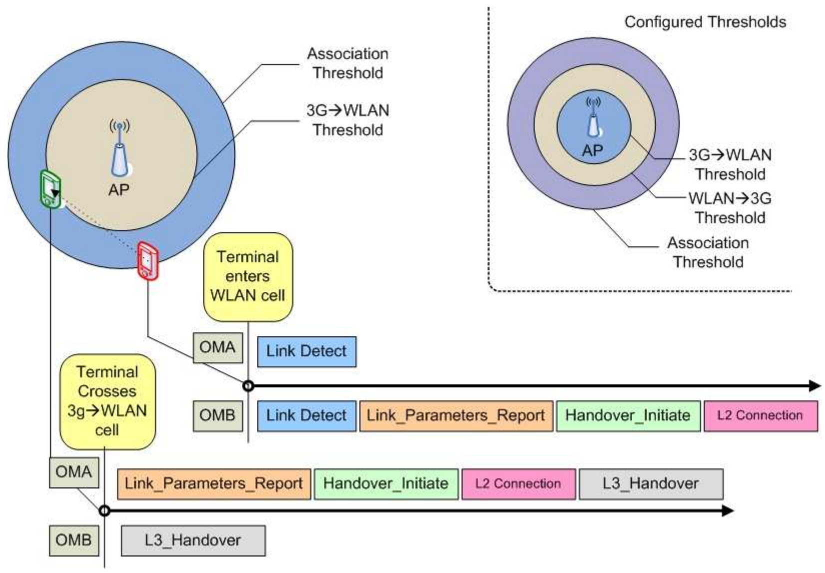

Figure 5.11: Different signalling stages for both operational modes

In OM.A, after cell detection, the terminal only triggers IEEE 802.21 signaling related to the L2 connection, i.e. message (2), at the 3G⇒WLAN threshold. So, in this mode, the terminal has to move within reach of a WLAN cell and the signal has to cross a certain configured threshold in order to execute a L2 connection, which is followed by the MIP binding update process. The nature of these operations ensures that good signal level conditions are met before the handover is executed, avoiding unnecessary handovers.

In OM.B the L2 connection related signaling is promptly sent when the association threshold is crossed, resulting in an earlier L2 connection. So, in this mode, when the signal strength crosses the 3G⇒WLAN threshold, the L2 connection has already been executed and the MIP binding update process can start right away. Note that the traffic flows through the old link until the L3 handover is completed.

For both modes, in order to maintain the handover's feasibility, upon receiving indication from the network to commit to the handover, the terminal's intelligence executes an active scan of the wireless environment. This procedure, executed in message (7) of figure 5.1, guarantees the signal strength hasn't deteriorated while waiting for the network handover command.

5.5.2.1 Simulation setup

The simulated scenario for operational mode differentiation shares the same characteristics of the one defined in section 5.2: an environment with a partial area of non-overlapping WLAN cells and full coverage of 3G technology. Each simulation was run with 3G⇒WLAN and WLAN⇒3G thresholds varying between -75dBm and -65dBm, as well.

Also, the same metrics where evaluated: Mean percentage of L2 connections without MIP registration, Mean number of 3G⇒WLAN handovers,Mean number of WLAN⇒3G handovers and Mean wireless utilization time.

5.5.2.2 Results

In this section OM.A and OM.B are compared against the above mentioned metrics.

Figure 5.12 represents the percentage of L2 connections not followed by a L3 Mobile IP registration (thus failed handovers), for both operational modes. The 3G⇒WLAN threshold variation from -75 up to -65 dBm shows us that, for 10m/s, the percentage of L2 associations not followed by a successful L3 handover increases up to almost 80% and 60%, for OM.A and OM.B respectively. The curves follow a similar increasing behavior for 2 and 5 m/s, although not so accentuated. As can be noted, the curves show a trend to increase while the 3G⇒WLAN threshold value is increased.

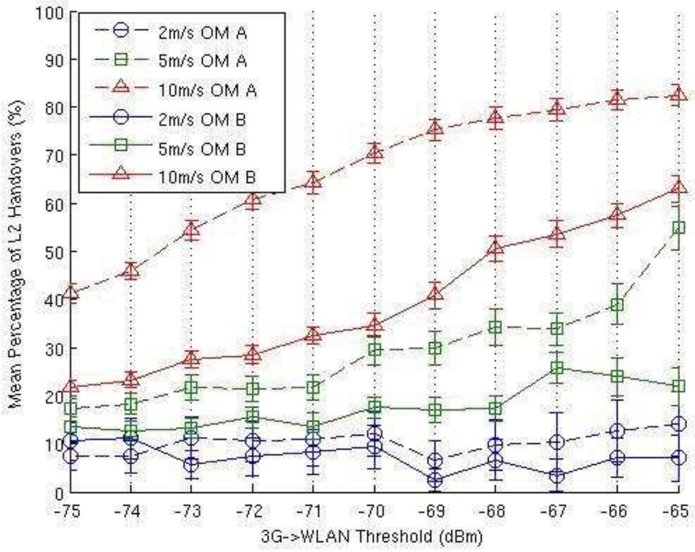

Figure 5.12: Mean percentage of L2 connections not followed by a L3 handover when WLAN⇒3G threshold is configured at -75 dBm

This is a direct consequence of the implemented signaling. As the 3G⇒WLAN threshold is increased (in dBm) the MN needs to be nearer to the AP to connect to the WLAN. Since the signaling handshakes that must occur are also impacted by the Round Trip Time (RTT) of the links, this interchange of information increases the probability of moving out of the cell prior to the reception of the MIH_Handover_Commit.request command. Depending on the speed of the terminal, this behavior can be dominant, as in the 10 m/s case. The different results obtained for the two operational modes shows that starting the 802.21 handover related signaling as soon as a WLAN cell is detected (i.e. OM.B) decreases the number of failed L3 handovers in 20%, for 10m/s. This decrease is also verified for the 2m/s and 5m/s speeds, although not as accentuated. This behavior is particularly evident at lower threshold configurations, where the percentage of failures is almost reduced by half. Figure

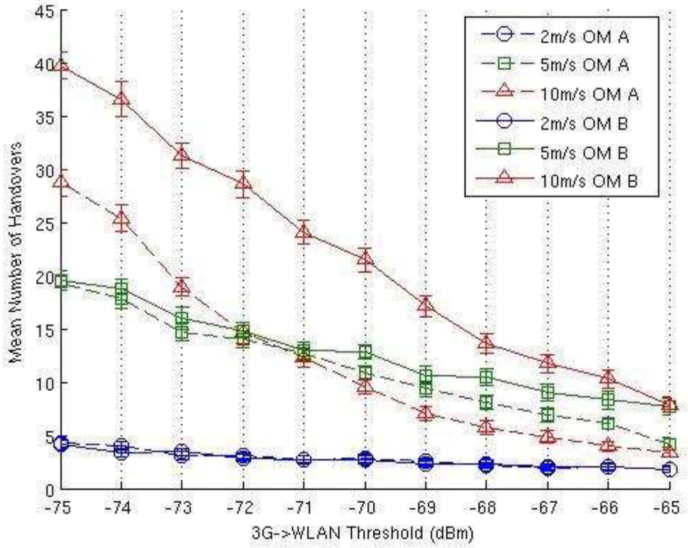

Figure 5.13: Mean number of 3G⇒WLAN handovers when the WLAN⇒3G threshold is configured at -75 dBm

5.13 depicts the mean number of L3 handovers obtained by varying the 3G⇒WLAN threshold. The impact of the speed affects the metric in different ways depending on the considered configuration. At the value -75 dBm the number of handovers is quite large especially considering high mobility level, while decreases and converges for greater values of the threshold. The decay in the slope of the different speeds is related with the failures of performing the L3 handover shown in the previous metric. The graph shows how the values tend to converge, when the 3G⇒WLAN threshold is increased.

It is interesting to note that the closer the mobile node to the access point, the lower the chance of having complete handovers. This is complementary to the previous graph in the sense that even starting the signaling at -80 dBm, thresholds close to the -65 value affect the feasibility of the handover procedure.

Comparing both operational modes, it is visible that the number of handovers in OM.B is greater than in OM.A, especially at higher speeds. It is also visible that at higher speeds the number of handovers decreases greatly, where the 10m/s curve becomes lower than the respective 5m/s curve, for OM.A. This shows how terminal speed impacts network response time making handovers not possible. In OM.B the 10m/s curve always resides above the 5m/s curve indicating that, by executing the signaling upon cell detection, there is a greater chance that the terminal is still inside the cell when the answer arrives from the network. Simulation results confirm therefore that, for a better and cleaner protocol design, splitting of the signaling for L2 and L3 events is required, where the L2 connection is performed at the association threshold and the L3 handover at the 3G⇒WLAN threshold. As a final remark on the tendency of the 5m/s slope and 10m/s slope we see how OM B outperforms OM A (when crossing the -72 dBm 3G⇒WLAN threshold) , reducing therefore the number of handover opportunities.

Table 5.5 shows the mean wireless utilization time for the different speeds, according to both operational modes. Variations on the 3G⇒WLAN threshold showed a flat behavior for all speeds, and the average value is represented here. As we can see, although OM.B has a slightly higher wireless utilization time on all three speeds, the point of start for the IEEE 802.21 handover related signaling has no significant impact on this metric.

	Operational Mode	
Speed	OM.A	OM.B
2m/s	32.25s	32.35s
5m/s	9.11s	9.65s
10m/s	4.33s	4.53s

Table 5.5: Wireless utilization time per handover.

The fact that within the same speed, the results for both operational modes are of the same order of magnitude also confirms that, when the handover is executed, performance of the configured system is maintained.

5.5.3 Access Point transmission power impact

When considering threshold configuration and time sensitive operations, it would be desirable to implement a model able to adapt to different environments such as operator dependent network deployments (e.g. network planning). This section analyzes the impact of transmission power on the threshold based model. A reference WLAN coverage area was chosen from the previous simulations, and new threshold values where calculated with new transmission values, maintaining the same WLAN coverage area for event triggering. The goal is to verify the model adaptation to transmission power changes, and to analyze handover behavior differences while maintaining the same signaling triggering points. Transmission power values were taken from commercial products data sheets, complying with UMA [53] and [58]. From figures 5.14, 5.15, 5.16 it can be derived that the previous metrics present similar values, within their respective confidence intervals, when comparing different transmission powers in the same model and at the same speed. Figure 5 shows how the first metric, incomplete handovers, maintains a relative linear behavior in which is visible the differentiation between OM.A and OM.B, particularly at higher speeds. The same linear behavior can be noticed in figure 5.15, for the number of handovers. Figure 5.16, for the wireless utilization time, shows us that there is no different behavior in handover related issues caused by transmission power or speed change. These results lead to the conclusion that changing the APs transmission power, at the same time than the thresholds, and thus maintaining the wireless coverage area, produces no changes in the results obtained.

Having this data in consideration, it can be argued that the dominant factor contributing for the result differentiation is the amount of time the terminal logically resides in the WLAN cell, which is affected by threshold configuration. It is desirable that this configuration can be dynamically adapted in the MN, through information received from the AP concerning its transmission value: this may allow an operator to dynamically adjust the coverage area of WLAN hotspots according to the density of users, but still retaining a uniform behavior in terms of handovers. This will allow operators to better configure their networks and terminals to improve handover procedures.

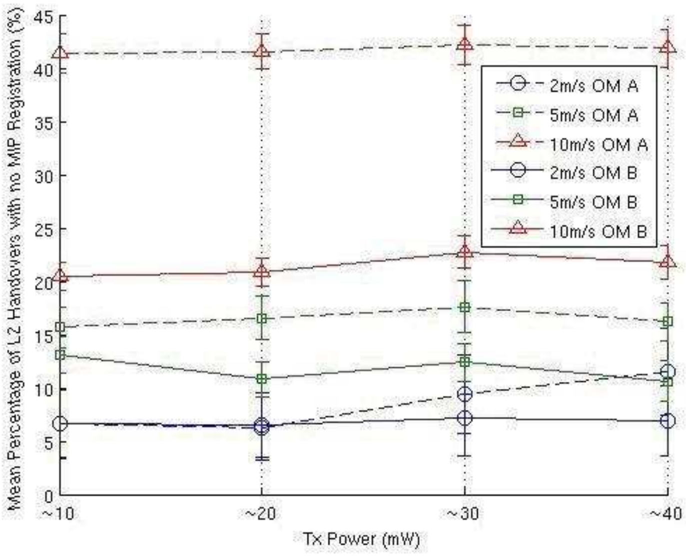

Figure 5.14: Mean percentage of L2 connections not followed by a L3 handover with different Tx and threshold values

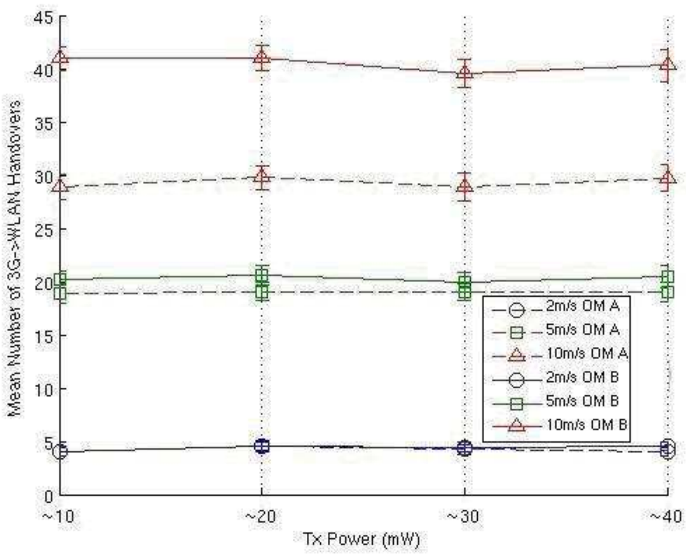

Figure 5.15: Mean number of 3G⇒WLAN handovers with different Tx and threshold values

5.6 Conclusions

This chapter presented the major contribution of this thesis showing that network driven mobility is achievable and that performance is acceptable even with real time applications such as Voice over IP applications. Extensions where done to the code of a simulation framework that enabled new studies to be done. The simulative study, starting from the IEEE 802.21 standard, specifies the design for both mobile terminal side and network side where the intelligence for enhanced handover decision making

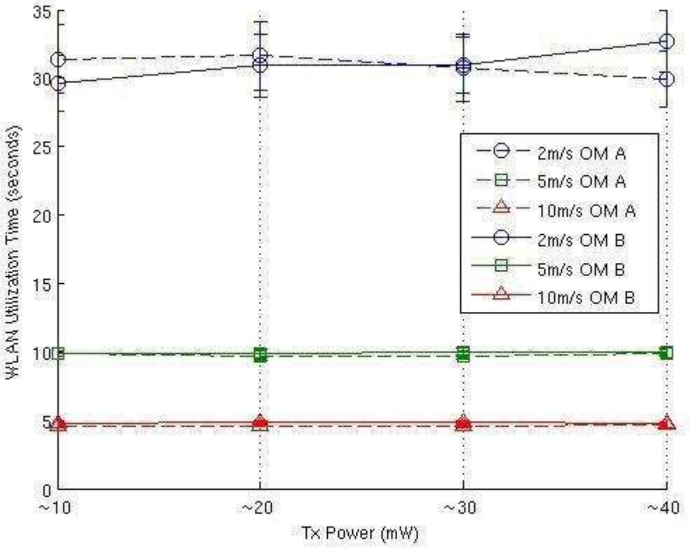

Figure 5.16: Wireless utilization time with different Tx and threshold values

is implemented. That is, while the terminal is able to trigger MIHO for poor radio conditions the network can perform admission control as well as load balancing by moving the mobile device from/to the WLAN/3G link. The approach here evaluated is IP based in a mixed 3G/WLAN scenario but, without loss of generality, also applicable to any other 802 family of technology, such as WiMAX, due to the abstraction facilities supplied by the 802.21 standard. Initially, a demonstration of the validity of the solution for indoor scenarios was done, followed by an analysis on how the system behaves in scenarios with a high degree of mobility. That is, the simulation proves that the signaling supports mobile users up to 30m/s, targeting scenarios such as WiMAX mobile scenarios. Such considerations can be relevant for the design of future 4G networks. In section 5.1.4 it is also demonstrated that the required bandwidth for handover signalling is at the most 300Bytes/sec when the MN moves at the speed of 10m/s. This is a negligible amount of traffic worth to be supported for network controlled mobility.

Finally, in section 5.5 it was shown how the designed model is radio signal level agnostic and the handover performance is only impacted from the time the terminal resides in the wireless cell. In fact, by adapting the thresholds configuration to the transmission power of the access points the wireless utilization time is not affected.

Chapter 6

Applying 802.21 to 4G Architectures

802.21 was introduced in a 4G network. The IEEE802.21 standard plays a part in the DAIDALOS project. The abstraction capabilities of the MIHF, hiding the underlying technology of nodes and PoAs, enabled core decision entities to use a common framework that simplified network operations. Also, services existing in the MIHF, such as the Event, Command and Information services, allowed, in one hand, to supply the means for mobile node entities to report current link conditions to the network decision points, as well as new links detects and, on the other hand, for the network entities to execute control over those mobile terminals.

The 802.21 MIHF as a technology independent abstraction layer, its services and mechanisms where successfully introduced in several publications. This section briefly overviews the 802.21 mechanisms which where applied to the DAIDALOS framework.

6.1 DAIDALOS Mobility Entities and Requirements

This section presents a description of the major entities involved in the DAIDALOS mobility process.

6.1.1 DAIDALOS Mobility Entities

Under the DAIDALOS architecture mobility mechanisms supporting seamless mobility across heterogeneous networks are coupled with QoS provisioning mechanisms [34]. QoS control is provided in a hierarchical fashion, where end-to-end QoS is separated from link-local QoS control at layer two. Layer 3 QoS inside the LMDs is managed by Zone QoS-Brokers (ZQoSB). These entities perform a wide range of functions such as per-flow admission control and resource management, handover authorization based on user profiles and available resources. They also consider resource optimization operations with respect to to both user requirements and network availability. An Authentication, Authorization, Accounting, Auditing and Charging (A4C) server manages user accounts and interfaces with QoS modules for authorization purposes as well as policies. This server is queried by other network decision entities when information regarding user profiles is required. User identities, service descriptions, contracts between users and service providers, roles and permissions, are examples of information stored in the A4C server database. Both MTs and the access routers, here denoted as MAGs-AR, contain elements that perform the enforcement of the QoS in the network and trigger the QoS process for admission control and resource reservation: QoS Client (QoSC) and QoS Manager (QoSM), respectively in the MTs and MAGs-AR. The entity responsible for mobility management on the terminal side needs to interface with the QoS modules (QoSC in this case) to provide seamless integration of QoS and mobility. In this sense, when mobility occurs, the mobility management needs to inform the QoSC to trigger the QoS reservation process. This integration needs also to be in place to support multihoming and the choice of interfaces according to QoS requirements. Also, it has the ability to evaluate either or not to accept a suggestion of handover made by the network side. The specific characteristics, and reservation handling, of each technology are executed by a Radio Access Layer (RAL), which is connected to a MIHF and is able to receive 802.21 commands and supply link-layer events. It should be noted that mobility related signalling between network entities residing in the core network is executed through the DIAMETER protocol.

6.1.2 DAIDALOS Mobility Requirements

When considering handover scenarios, it is crucial to preserve QoS. Although L3 QoS is extremely important, preserving L2 QoS may be more critical since, depending on the technology in place, it may require some time to request that QoS level from the technology, creating a noticeable impact in the user experience. The details presented in this section are more concerned with the L2 QoS case. The conceived solution in DAIDALOS to maintain QoS during handovers is to split the handover procedure in two stages: preparation and execution. During the first, involved APs are notified for L2 QoS resources preparation in advance, so that when the handover is carried out they simply activate those previously prepared reservations. For the execution phase, the upper layers may need, for some technologies, to notify the RALs of the exact time when handover is finished. In order to establish L2 QoS resource reservations, the QoSM requires topological knowledge of which L2 entities are on path to the mobile node, defining to where resource preparation requests should be sent. Also, for this matter, it is required a mean to supply link commands, with origin at high-level entities, to several different technology link layers. Lastly, these mechanisms need to be integrable with L2 and L3 mobility mechanisms.

These requirements positioned IEEE 802.21 as a prime candidate to supply these functionalities, whose application to DAIDALOS is discussed next.

6.2 802.21 Mechanisms Supporting DAIDALOS Architecture and Interfaces

802.21 was adapted as the control plane protocol, covering both vertical and horizontal handovers. This allowed MIH mechanisms to complement data plane control protocols, these being MIPv6 (for GMD) and NeTLMM (for LMD).

DAIDALOS considers a MIHF at every entity (either network entity or mobile terminal), allowing for full exploration of all available 802.21 mechanisms considering the supply of information to enhance the handover decision.

The following figure supplies an overall view of the network architecture.

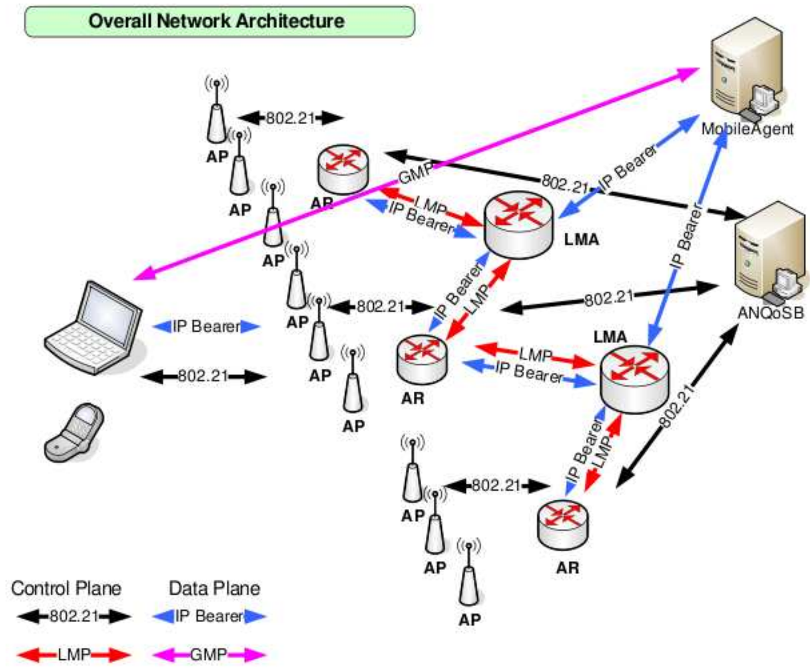

Figure 6.1: Network Architecture

This figure allows us to see how 802.21 communication between all entities that can support Local and Global Mobility procedures, where events generated at the mobile node (and also the Access Points), reach the network decision point (in this case composed of a Mobile Agent and a Access Network QoS Broker) allowing them to be aware of the terminals (and PoAs) point-of-view of the network.

An interesting aspect of using 802.21 under this framework is that, where in the picture is an AP, it could be another technology Point of Access: abstraction services provided by the MIHF allow the MIH-users (able to take handover decisions) to supply a general command to the MIHF which then translates it into primitives of the required technology.

6.2.1 MIH-enabled Entities Discovery

In order to make available its services, 802.21 requires that MIH-enabled entities advertise and detect MIH capabilities of the nodes around each other. This was considered from the beginning in DAIDALOS, allowing network nodes to acquire topological knowledge of its neighbors (even inferring a inter-technology map using the MING described in section 3.4), as well as to acquire knowledge of which services, events and commands are supported.

6.2.2 Events Reporting

Typically, in IP networks, mobile terminals have the ability to decide when a handover has to be performed in order to maintain ongoing sessions in the most seamless manner possible. In scenarios where several heterogeneous networks are involved, it is not realistic to keep all information about existing networks in the mobile terminal because this would imply a higher degree of complexity in terminals and create a significant amount of signalling overhead in the access network, to support that complexity. This means that information and also handover decision algorithms can be located, centrally or distributed, all over the network.

DAIDALOS architecture has been designed in order to allow Network Initiated Handovers (NIHO) in addition to Mobile Terminal Handovers (MIHO). In this architecture, a network entity decides, based on all information available when, how and where to mobile terminals should execute handovers. In order to take such a decision, the decision entity needs accurate and up-to-date information. This information can be generally classified into static and dynamic information. Static information, such as cost, QoS capabilities or mobility protocols supported by certain networks, can be obtained upon request to an information database. However, dynamic information, such as load in certain network nodes or networks perceived by the terminal, has to be obtained through more actualized and accurate means, based on event triggering. Events can be triggered either by end mobile terminals or by other network nodes involved in the communication path.

802.21 supplies the required mechanisms to support such information exchange. Mobility Management Entities residing at core network entities can register at the terminals MIHF, and subscribe for the collection of evens pertaining to radio link conditions (i.e., indicating deterioration of signal) as well as detecting available new links. This information can serve as input to the network to initiate handover procedures for optimizing the user experience or managing load amongst its access networks. Also, Mobility Management Entities can reside at the terminal and are able to also collect these events using exactly the same mechanisms that are available to the network entities, thus favouring a simplistic common approach for either MIHO and NIHO.

6.2.3 Command Execution

Upon detection and registration on the mobile terminals MIHF, MIH-users residing at network entities (as well as in the terminal itself) are able to collect indications of current conditions of the terminal. Network entities can even further combine these inputs with events from network entities, and execute handover actions.

These commands can also be used to obtain information of a more static nature (such as MAC addresses, type of available technologies, etc.) to even further enhance a handover decision.

6.2.4 Quality of Service

Terminal bootstrap (i.e., first time activation of the terminal) or handover cannot be executed without Quality of Service guarantees.

As explained in section 2.4.3 the DAIDALOS framework encompasses several local domains (or Access Networks) connected to a core network. In order to provide optimal servicing for users in these networks, in a broad sense, it is required to establish reservations in the links transversing the data path from and to the terminals. These reservations have to consider parameters such as bandwidth, throughput, jitter, etc. Thus, a common framework needs to provide the means to enforce such reservations in the specified links, as well as to detect when and where that enforcement is required. 802.21, in DAIDALOS, also had a part here, albeit requiring some extensions which are described in section 6.3.2.

MIH-enabled network entities could use 802.21 signalling to enforce such reservations upon network entities residing at the access network. Moreover, the MING could be used to inform (at specific nodes) topology information, allowing that specific network entity to know which other network entities where in the datapath towards the terminal, and thus to where should reservation commands be sent. In this way, core entities are relieved of having to know complete topology knowledge of Access networks, delegating that function into "border" routers.

Also, using 802.21 signalling to enforce reservations allows to seamlessly integrate the reservation process into the normal 802.21 handover signalling, studied in Chapter 5.

6.3 DAIDALOS Extensions to 802.21

Some extensions to 802.21 were required in order to support DAIDALOS requirements. These extensions consider new mechanisms such as having an entity that is able to proxy 802.21 messages between the terminals and the broker, the introduction of Quality of Service primitives and the design of a new procedure for handover control.

6.3.1 Proxy Action

As was indicated previously, the DAIDALOS architecture separates the core and access networks. In order for a 802.21-enabled node to register itself in the network and to use its services, it has to connect to a network node from which it can obtain those services. This places the core nodes as prime candidates for this connection, since they are aware of information such as Quality of Service conditions throughout the network. However, for a mobile node to have direct access to core network entities can pose a serious threat, generating security and privacy issues.

For that matter, a MIH Proxy behavior was envisioned. Under this assumption, the MN would not connect directly to a MIH-enabled entity in the core, but would do it so to a network node in its Access Network. This node would then "proxy" 802.21 signalling into the core, protecting core nodes from tampering, and allowing functionalities such as:

- Implementation of different authentication security measures (802.1X, KERBEROS, etc.) on each Access Network depending on issues such as technology, operator, etc.;
- Implementation of other transport protocols (RADIUS, DIAMETER, NSIS, etc.) in the core, different from the ones used in the Access Network.

6.3.2 Quality of Service Primitives

The extended MIH primitives for QoS where MIH_Handover_Prepare.request and MIH_Handover_-Prepare.response. These are 802.21 messages exchanged between MIHFs of network entities and, in their normal behavior, are used to check for QoS availability in those entities. The message indicates parameters to be analyzed at the target:

- Throughput;
- Packet Loss Rate;
- Packet Error Rate;
- Class of Service Minimum Packet Transfer Delay;
- Class of Service Average Packet Transfer Delay;
- Class of Service Maximum Packet Transfer Delay;

- Class of Service Packet Transfer Delay Jitter.

In normal 802.21 behavior, these parameters are used as threshold definition points that, upon crossing, generate events which can then trigger handovers.

The envisioned behavior for DAIDALOS differed from this, however. The parameters supplied in this primitive where also used to implement the reservations at the specific links. In this manner, the MIHF would then translate those commands into link specific primitives able to implement the required reservation.

6.3.3 Network-Assisted Mobile Initiated Handovers

802.21 allows a clear definition of who is the initiator of a handover, regarding Network Initiated Handovers and Mobile Initiated Handovers. In the first case, a MIH_Handover_Initiate.request message is sent by a network decision entity to a mobile node, and in the second case this order is inverted.

From the operator point of view it is architecturally complex to implement a common approach supporting MIHO and NIHO at the same time. The causes for this pertains to the fact that both MIHO and NIHO operations use the same messages and parameters, but being sent to different entities under different contexts. So, it is not possible to infer a generalized approach considering both NIHO and MIHO, if separation of the mechanisms is done.

However, in DAIDALOS, an integration approach has been researched: Network-Assisted Mobile Initiated Handovers (NAMIHO). The aim of this approach is to supply all the mechanisms from both NIHO and MIHO, but under a common signalling. Under this concept, the MIH_Handover_-Initiate.request message is always sent by the terminal to the network. But, a new command, MIH_-IIS_Trigger_Command.request can be sent previously by the network to the terminal. This command is used for "emulating" a NIHO, where the network, after receiving events from the terminal, deems a handover is necessary. So, this new command is sent towards the mobile node, with a list of possible handover candidates. This list is further processed by the mobile node, which then undergoes the handover process by sending the normal MIH_Handover_Initiate.request message to the network. The MIHO behavior can be achieved when the terminal, via detecting and choosing a new (or a list of) new network, just sends the MIH_Handover_Initiate.request message to the network.

6.4 A DAIDALOS Handover Scenario Example

In the next figure a handover scenario is considered where two 802.11 networks are directly connected through a core network and are managed by the same ZQoSB.

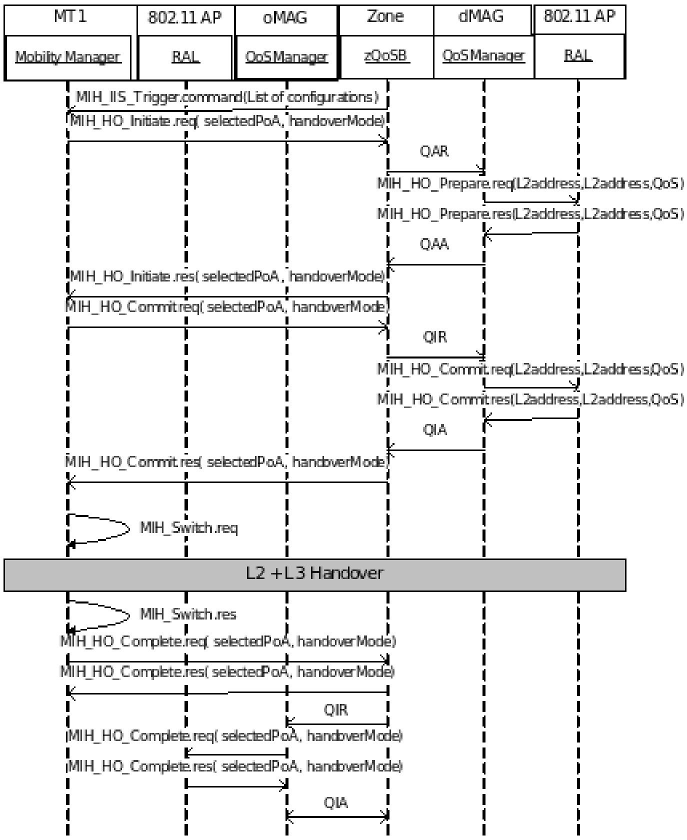

Figure 6.2: DAIDALOS Handover Signalling with QoS Example

The process starts with the ZQoSB supplying a list of possible handover configurations to the MT, through a MIH_IIS_Trigger.command. This list is processed at the terminal which selects a feasible handover target and issues a MIH_Handover_Initiate.req command to the ZQoSB. This entity acknowledges the possibility of handover to that target AP, and triggers the reservation process, sending a Diameter QAR message to the QoSM of the nMAG that connects to the AP. This message reserves L3 resources in the new network and triggers L2 resources pre-reservation on the required network entities in the path to the target AP, sending them a MIH_Handover_Prepare.req command. In this case, a 802.11 AP receives this command. A response, containing the pre-reservation results, is sent all the way back to the terminal, which can now commit to the handover, supplying a MIH_Handover_Commit.req to the ZQoSB. A Diameter QAA is sent to the new network activating the previously reserved resources, triggering a commit command to execute the activation at link level. A response containing the result of this activation is sent all the way back to the terminal, with an indication that resources at the handover target are prepared. At this time the handover can take place, and a respective switch command is sent by the terminals Mobility Manager to the respective RAL (through the MIH layer), which will execute the handover process. At the end of this process, a MIH_Handover_Complete.req is sent to the zQoSB, acknowledging it that the handover has finished successfully. This entity also acknowledges that the resources at the old PoA need to be released. For that, it sends a QIR message to the QoSM at the oMAG, triggering a MIH_Handover_Complete.req to all network entities involved in the data flow of the previous PoA of the terminal, clearing the no-longer used resources. When this process is finished, a response with the results of this operation is sent to back to the QoSM at the oMAG.

6.5 Exploration in the IETF of 802.21 concepts developed in DAIDALOS

With the work done in DAIDALOS it was verified that the inclusion of a layer 2.5, such as the MIHF, could provide mechanisms (such as media independent link layer events and commands) improving the MN-AR interface under a common framework. In the IETF, this issue was also addressed in recent proposals for network-based localized mobility management [28] [32] [33].

This subject was explored and a Problem Statement was derived in a Internet Draft [6] submitted to the IETF and presented at the *68th IETF Meeting* in Prague, in March. This section provides an overview on the draft.

6.5.1 Issues

One important issue pertaining to the PMIPv6 concept is that no special host modifications should be done. The usage of a common 2.5 layer such as the MIHF can provide, on one hand, the means to apply link layer information required for node attachment to the network and, on the other, the possibility for mobility management entities in the network to control those link layers in a remote technology independent way. Other approaches for MN-AR interfaces exist such as [59], however this approach relies on Layer 3 connectivity for link change detection, contributing for a more time consuming mechanism.

Also, considering PMIPv6, current trends point to the use of the MAC address as the mobile node identifier. However, this is restrictive in a way that the mobile node would be restricted to the L2 address of a single interface. Using identifiers provided by the MIHF, it is then possible to address multiple interfaces, in a technology independent way, through the same identifier.

The 802.21 standard also provides the means to exchange static information through the Media Independent Information Service, allowing for context information exchange between PMIP entities, when required (for example, to supply MN identifying information to a handover target Point of Access).

6.5.2 Identified Scenarios

Under this study several scenarios where identified. First, when the terminal powers on (e.g. Boot-up), the host must be able to provide link layer events and indication in case a successful attachment occurs. This allows for PMIP-aware network nodes to acquire information regarding the node. Second, proactive handover scenarios require that the terminal supplies information regarding the detected handover target and current radio conditions. This information can be sent using the Media Independent Event Service of the 802.21 standard. Lastly, in reactive handovers the terminal can supply a link layer event indicating successful attachment.

6.5.3 Requirements

Being a Problem Statement, this document does not intend to provide (at least for now) solutions for the issues encountered. However, some requirements where specified which should be taken into account.

- The framework should be able to be incorporated with future instances of the NetLMM protocol, as well as optimizations of it;
- The framework intends to enhance and improve L3 procedures and not to replace them, facilitating the information exchange between the MN and AR;
- A clear communication model between between the MIHF and the upper layers needs to be defined;
- The MIHF identifiers can be used under this framework and should be aligned in a common framework with the remaining identifiers (e.g., L2 addresses);
- Event registration from network elements to the MIHF at the terminal node has to be done carefully, since events being sent to multiple network management entities can trigger unwanted parallel behavior;
- The host must be able to store previous attachment information;
- Security considerations, such as an attacker node supplying false link layer events, have to be taken into account.

6.5.4 Future Work

The work was presented to the MIPSHOP WG raising some interest from some of the participants. Ongoing work is being done in terms of refining the draft, and providing some technical appendixes referring to specific considerations on the usage of the 802.21 standard. If deemed necessary, there might also be the chance to propose new extensions to the IEEE 802.21 standard, based on the results obtained with the IETF drafts.

Chapter 7

Conclusions

7.1 Conclusion

This thesis presented concepts and issues pertaining to heterogeneous network controlled handovers. In chapter 2 it first introduced the user to wireless technologies like 802.11, 802.16 and 3G cellular networks. Then, it evolved into the definition of multimode terminals where these technologies can be combined into a single device. The infrastructural means that define networks able to support these approaches are presented in section 2.3, showing the current trends in mobility, MIHO and NIHO, and in 2.4, where the DAIDALOS project was presented.

These sections present the requirements, issues and goals for 4G networks, which served as the motivation for approaching the IEEE 802.21 standard. After studying the MIH entities, services and procedures in chapter 3, an extensive study of ongoing 802.21 activity was done. Moreover, a simulation implementation where a MIH-enabled node executed handovers between 3G and WLAN was studied. This study, however, focused on Mobile Controlled Handovers where the network had no involvement regarding L2 handovers. This was the startup point for an extensive implementation effort to extend the simulator to support terminal and network mechanisms which allowed the study of network controlled handover concepts. The algorithm of the mobile node was implemented as in figure 5.2 and a new network entity was introduced for centralized support of the signalling depicted in figure 5.1.

This enhanced simulator allowed to show results in section 5.2 where the benefit of applying network controlled handovers in scenarios where typically mobility decisions are host centric. This also allowed to conclude, in section 5.6, that NIHO is suitable for next generation 4G networks. More concretely, it was verified that the 802.21 mechanisms allow link layers to support measurements and information about the radio conditions that they are experiencing, to high-level entity modules that are responsible for mobility management. Also, 802.21 allows this information to be conveyed to remote entities, enabling network entities to receive indications pertaining to cell detection and radio conditions experienced by the terminal. This information can be used as an input for mobility decision engines located in the network, which are themselves able to use 802.21 commands to send handover orders towards terminals. The signalling between remote entities was also studied and verified to pose minimum impact on the performance of handovers.

The study presented here can be seen as an useful beacon for next generation mobile operator networks, where new methods are required for better support of mobile subscribers across heterogeneous networks. The authors of [60] highlight new possibilities for the optimal delivery of IP based services, considering NIHO as one of the possible solutions. With the work done in this thesis, and the publications that followed, the IEEE 802.21 technology was successful presented as an enabler of both MIHO and NIHO solutions, and also supply considerations on its applicability on 4G network environments as well as new trends in mobility schemes.

On a more personal note, the work developed throughout this thesis enabled the author to work on a very interesting topic, the 802.21 standard. On one hand, it was possible to study and analyze a new and upcoming technology that is still under active discussion and, on the other hand, it enabled a new way of thinking that gave birth to the analysis of several new subjects and architectures, and how 802.21 could be applied and achieve a prominent role. The need for performance analysis of 802.21 mechanisms, as well as its application to handovers, lead to the use of network simulators to implement the required mechanisms and obtain results. This enabled to not only learn how to work with a powerful simulation tool such as OMNeT++ and to investigate an existing implementation of 802.21 mechanisms on that simulator, but also allowed the development of statistical studies for result analysis. Later on, placing emphasis on network-controlled handovers, new code was built and added to the simulation framework which allowed the observation of new behavior and the production of new results, which were published in a conference paper. These results, and the extensive work developed about handover mechanisms, allowed the author to apply this knowledge into an integrated project of the European Union, the Daidalos project. The work done here was also very interesting, empowering the author with new concepts regarding complex 4G architecture mechanisms that supported an advanced new communication infrastructure for the future. Here, the application of 802.21 mechanism took a prominent role, and the author was actively involved in discussions and the design of solutions for its use. During the course of dealing with these architectural issues, the author was also involved in standardization effort with the IETF, related to the application of 802.21 to new trends in localized mobility.

Finally, regarding the work done for this thesis, the chance to work on an exciting new technology and its application to a project that aims for the betterment of the technology that will be among us in a near tomorrow, constitutes an important mark not only for this thesis, but also to the author's own life.

7.2 Future Work

The IEEE 802.21 standard is an on-going effort and is still under conception. Investigation work able to support new contributions to the 802.21 standards, as well as on networking conferences, is always possible and welcomed.

Currently, the effort on IETF to introduce 802.21 mechanisms to several areas as lead to the creation of a Internet Draft [6] and its on-going support and extension. This will also be subject to publications, providing insight, and possibly results, on the proposed mechanisms.

802.21 is being introduced to key areas that were not thought of during its initial definition. An interesting example is the integration of Media Independent mechanisms with pre-authentication methods, enabling scenarios where mobility optimization can be coupled with the secure crossing of mobility domains. An example of this approach can be found in Media-independent Pre-Authentication (MPA) mechanisms [61] that have started to appear on IETF charters.

Another point for future work is the comparison of the obtained simulation results with new simulation frameworks that have become available recently. An example is the *The National Institute of Standards and Technology* (NIST) simulation framework[1], done using the *Network Simulator* (ns-2)[2],

[1] http://www.antd.nist.gov/seamlessandsecure/toolsuite.html
[2] http://www.isi.edu/nsnam/ns/

that implements support for Media Independent Handovers (IEEE 802.21) in IEEE802.16e technologies. On one hand, it would be interesting to work with a different simulator than OMNeT++, and on the other hand, results involving a new link technology would be interesting to analyze.

Simulation work can potentiate interesting scenarios that can prove difficult to access in real-life platforms. Examples of these scenarios can be the introduction of large numbers of user terminals. However, considering a real 802.21 implementation is always an interesting challenge, considering the architectural issues involved and the verification of simulation data in real-life scenarios. A real implementation would consider solving not only network related issues, but also technology and computational ones, constituting an interesting problem that would have to be tackled in different fronts.

As always, mobility is always a hot topic: users always on the move, terminals with increased capacity, new and more demanding multimedia applications... these are all factors that push the technology edge to its limit, demanding an on-going effort for improvement on all fronts. New and futuristic mobility scenarios are always around the corner. Analysing (and even designing) the part of 802.21 in all this is a task that will generate ample ramifications of future work.

Appendix A

OMNeT++

OMNeT++ is a object-oriented modular discrete event network simulator. An OMNeT++ model consists of a hierarchy of grouped modules. The depth of this group of modules is not limited, allowing the user to define the logical structure of the simulating systems, through a module structure. Modules communicate between resolved trough message passing, that can contain simple of complex data structures. Modules can sent messages directly to their destinies, or trough a pre-defined path, using gates and connections.

Modules can have their own parameters, which are used to customize their behaviour and topology. Modules places in the lowest hierarchical level are responsible for encapsulating the system's behavior. These modules are nominated Simple Modules, and are coded in C++ using the simulator library.

This simulator can be used for:

- Telecommunication network traffic modelling;
- Protocol modelling;
- Queue modelling;
- Multi-Processor and other distributed hardware systems modelling;
- Hardware architectures validation;
- Performance evaluation of complex software systems;
- ... modelling any kind of system where it is possible to do a discrete event approach.

OMNeT++ can be freely used for academic or non-profit ends.

A.1 Modelling

As was pointed before, modelling in OMNeT++ has its basis on the hierarchical modular structure, with message passing using channels. It also allows a flexible module parametrization, and contains its own topology description language, NED.

Modules that contain other modules are Compound Modules. Stand-alone Modules are Simple Modules, that reside in the lowest hierarchical level. All modules are instantiated as sub-modules of the system modules.

Connections between modules have three parameters that facilitate communication network modelling: i) propagation delay, ii) error rate and iii) data transmission rate. It is possible to define parameter individually per link, or to define link types and reuse them in the model. One has to regard an important detail of OMNeT++ related to data transmission rate. When sending a message in the model it is the same as sending the first bit. Receiving that message is receiving the last bit. This factor is important, for example, in modelling Token Ring networks, when the destination does not wait for the reception of the whole packet, but starts retransmitting bits as they arrive.

Regarding simple modules, which are coded in C++, all flexibility and power of the programming language can be used, since it is totally supported by OMNeT++ simulation classes. The simulation programmer can opt between an event-oriented approach or a process-oriented approach, and can freely use object-oriented concepts (inheritance, polymorphism, etc.) and design patters to expand the simulator's functionality.

Simulation objects, such as messages, modules, queues, etc., are represented by C++ classes. They where designed in a way to effectively interact, creating a powerful simulation framework. The following classes are part of the simulation library:

- Modules, Gates, Links, etc.;
- Parameters;
- Messages;
- Containers;
- Data collection classes;
- Statistical classes;

A OMNeT++ model is composed of the following parts:

- A topological description in NED language, pointing the module structure with parameters, gates, etc. This description is stores in .ned files.
- Module source-code, composed of .h and .cc files.

The simulation system provides the following components:

- Simulation kernel, containing the simulation generation code and the simulation class library;
- User interface, used for the simulation execution, debugging, demonstration our serial simulation execution.

Simulation programs are written from the above mentioned components. All source-code is compiled and linked with the simulation kernel and the user interface, forming a simulation executable. The .ned files are dynamically loaded in their text form at the start of the simulation program.

A.2 Graphical Interface and Results Output

The simulation environment can be executed in OMNeT++'s graphical interface. It allows to visualize the inner parts of the simulation model, execution control and, in some cases, alter variables inside of the model. This graphical visualisation is extremely important in the development phase and debugging of a simulation project, since it allows the programmer to obtain and analysis of the model behavior. Also, the graphical interface can also be used as a demonstration tool of a model's operation.

The output of a simulation is written in data files with vectors and scalars. OMNeT++ includes tools to visualize those files, Plove and Scalars.

A.3 IPv6SuiteWithINET

Separate frameworks built on top of OMNeT++ exist, which contain implementations of protocols, nodes, etc. These frameworks enrich the range of simulation options of the basic model. One of those frameworks is IPv6SuiteWithINET, which was used for this project. It supports IPv6 including models for MIPv6, PPP, Ethernet and 802.11. Also, it includes mobility support, allowing the definition of trajectory movement of the nodes.

A.4 OMNeT++ Mailing List Participation

During the course of this project, particularly in the OMNeT++ implementation part, several comments, suggestions and all-round support was done in the OMNeT++ Mailing list, when the opportunity and subject presented themselves.

Appendix B

Implementation Details

B.1 Extending the Terminal Architecture for NIHO support

The architecture of the terminal is based on the one from [56], but extended to support a subset of the Media Independent Handover Protocol defined in [1]. This protocol defines frame formats for exchanging messages between peer MIH Function entities, based on the primitives from the MIES, MICS and MIIS. The protocol also provides services for MIH capability discovery, MIH remote registration and MIH message exchange. As described in chapter 3, a MIH Function on a peer node has to register to the MIH of another node, in order to be able to receive events from there. The implemented functionalities focus on the impact of the signalling required to perform handovers, so MIH capability discovery and remote registration are supposed to already have occurred. The specification supports several transport options for the MIH Protocol, depending on the technologies and layer used. In this simulation implementation, it was opted for a standard L3 protocol using UDP, according to [41].

The terminal's handover algorithm described in [56] is still used, but with its behaviour preceded by the communication between the terminal and the PoS. The handover algorithm is required for the MIP Registration phase making the communication between the MN and the PoS mainly affect the L2 handover, where the MN only starts the handover if the respective command is received from the PoS. This command is sent upon reception of an indication from the MN about a new link, and favourable handover conditions. Extensive alterations to the OMNET++ Wireless Network Interface module were required, where Link Events were implemented in the MN, which are sent to the terminal's MIHF. These events are forwarded through a Media Dependent SAP between the Wireless interface and the MIHF, and are motivated by link detection and signal strength interpretations for threshold crossing. More details on the originated events and their subsequent transmission to the PoS are described in section B.2.

For the message exchange, a new UDP application had to be created in OMNET++ to process the remote messages received by the PoS. Also, since there is a video stream being sent to the terminal in parallel with the 802.21 signalling, the UDP layer had to suffer some alterations to process and distinguish both flows (namely, to avoid sequence number overwriting). A pre-defined port was defined for UDP communication. Respectively, a new UDP application had also to be written for the MN: in order to process the messages received from the PoS and sent by the MIH Function at the MN.

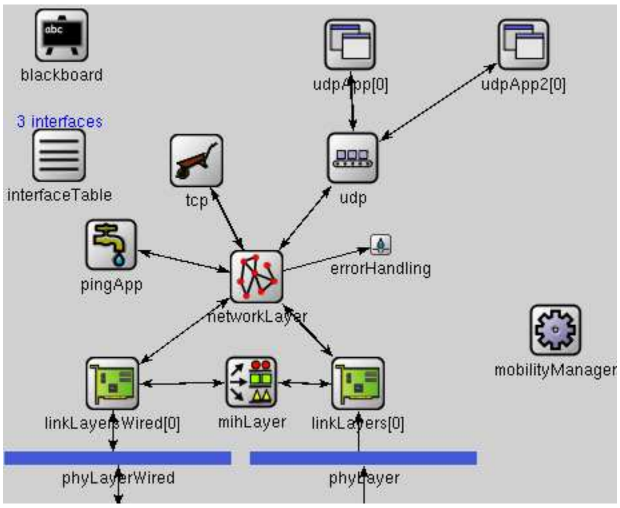

Figure B.1: MN module structure for the implemented NIHO framework

As can be seen from the picture, the new UDP application is connected to the UDP layer of the protocol stack of the MN. The UDP layer module encapsulates messages sent from the UDP application, and decapsulates messages received from the network layer. The MIH Function in the MN also communicates with the UDP application, supplying MIH events and MIH commands.

Regarding messages and signalling, in the case of the MN, messages originated at link levels had first to be sent and processed in the MIHF and then, if necessary, be relayed to the PoS' MIHF. At this point an important issue was found relating to the transport of .21 messages. What should be the terminal's behaviour in case a message is delayed or lost on the network? Should the problems raised through these issues be solved by the 802.21 or should they be handled by the underlying network? These transport issues are an important part of the MIH Problem Statement in [37], but it covers no concrete solution. To avoid the loss and collision of .21 signalling messages a network topology was adopted that relieved the strain from the communication link between the MN and the PoS. Also, the delay on the links belonging to the network was reduced to reflect a proper distance between Access Points, simulating for example, a network in a university campus or an airport.

Upon reception of a message from the MN, the PoS will react to it, considering if the load balancing is working or not. More details on the PoS behaviour are described in section B.2.

B.1.1 OMNeT++ Module Modifications

The introduction of MIH messages allows the MN to supply fresh information about link conditions to the PoS. The main behaviour of the new implemented content was introduced in the Wireless Ethernet Network Interface module, and the MIHF module for the MN created in [56]. In the first, several important processes are defined such as the Active Scanning procedure and the association to a previously found Access Point. In the later, the Media Independent Handover function for the MN is defined, comprised of a MIES service and a handover mechanism. Basically, the changes made to the Wireless Ethernet Network Interface module where in altering the normal flow of the wireless behaviour, prior to association of the MN to an AP. The changes done to the already implemented MIES where the processing of new link layer events, sending MIH signalling to the PoS through the new designed UDP application, and reacting to commands sent by the MIH Function in the PoS.

The association process is done through the normal procedure already existing in the OMNeT++

simulator, comprised of an Authentication phase and Association phase. The Authentication phase begins when the Wireless Ethernet Network Interface of the MN finishes an Active or Passive scanning, and finds a suitable AP, i.e. whose signal level received through Probe Responses is above the association threshold. In the normal model, this association was done as soon as the signal level crossed the Handover threshold, set in OMNeT++ at -80dBm by default. In the executed simulations, the state flow was altered to introduce the signalling between the MN and the PoS. Since the study focused on the effect of the signalling required for handovers through the use of thresholds, a new set of states are created (visible in figure 5.2), in which the signalling from the MN and the PoS is triggered at a defined 3G→WLAN threshold ranging between -80 to -65 dBM. When the signalling exchange is finished, and the necessary message for handover initiation is received from the PoS, the MN is able to resume the normal flow of the association process (L2 handover). This allows, for example, when the PoS determines that a 3G→WLAN should not occur, to prevent the L2 handover from being executed, and enforcing the PoS decisions. Note that, in this architecture, the process for association is not immediately resumed. During the course of time between sending the Link Parameters Report and receiving the Handover Commit Request, the signal could have changed and no longer present the necessary conditions for a handover. So, when the MN receives this message, a new Active Scan is made in order to: a) verify is the signal still presents the necessary conditions for handover and, b) in case of contemplating inter-AP handovers to verify if a better PoA target presents itself. Upon entering the Authentication phase, an authentication request is sent to the target PoA, the selected AP. When a positive authentication response is received, the MN enters the Association phase and sends an association request. When the MN receives the association response, the terminal is associated to the target PoA and starts listening for its beacons. In the OMNeT++ model, the beacons are sent with a frequency of 100ms. At this point, the MN is ready to send and receive data sent through the WLAN network. After this phase of connection to the Wireless network, the MIH Function at the MN is informed, through a Link Event, LINK UP, that the wireless interface is connected, and starts evaluating the signal level received through the beacons.

According to the handover mechanism defined in [56] upon favourable evaluation of the mean of this signal level, a L3 handover can occur, through the realization of the MIP Registration phase. In this phase, the MIP signalling and binding occurs, to indicate to the HA the new CoA of the MN. Due to the threshold being simulated, and also to the movement of the node and the delay cause by the signalling, a L2 handover might not lead to a MIP Registration. This is one of the metrics of the simulation model, which is extensively studied in section 5.3. Since this is a make-before-break handover, the MN will attempt to establish the new link before releasing the old one.

Another set of thresholds used in the simulations is the WLAN→3G thresholds, defined as {-80,-75,-70}. When the MN is receiving through the Wireless interface, its MIH also collects the signal level from the AP's beacons. Through this signal analysis, the MIH Function at the MN can verify if the signal level is still in favourable conditions or not. When the WLAN→3G threshold is crossed, the MN informs the PoS and this initiates a handover.

B.2 Network Functionalities

The PoS is a network entity whose MIHF is registered to the MN's own MIHF, receiving the subscribed events. Through the received messages, the PoS gains awareness of the terminal's position (regarding detection of WLAN cells) and the quality of its received signal strength. Then, the PoS can supply a remote command for handover initiation depending on the load value for that AP.

B.2.1 PoS Module Implementation

The PoS was implemented as a network node with the normal intrinsic modules (physical layer, link layer, network layer, UDP layer). A special UDP application was built on top of the UDP layer, with the ability to receive MIH signalling encapsulated in UDP packets sent by the MN, and also to send MIH signalling as well.

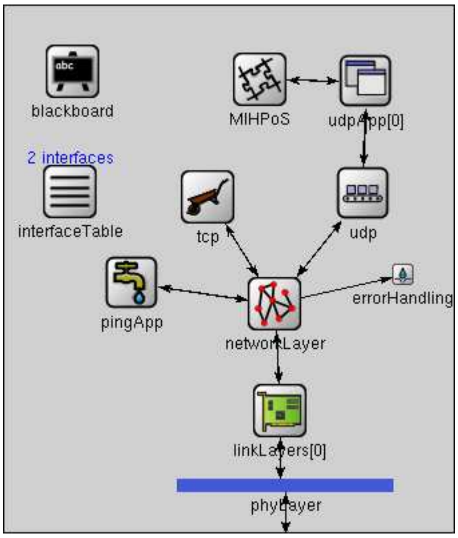

Figure B.2: Module Structure of the PoS

A new MIH Function was also developed and connected directly to this UDP application. On one side, this differs from the MN's implementation which has it's MIH Function connected at the link layers, since it deals with Link Events and Commands. On the other side, the PoS's MIH was planned only to deal with remote messages so there was no need to connect it to the PoS' Link Layer. Also, this reflects the nature of the MIH Function which, according to the latest version of the draft, is not only connected to the Link Layer and the Network Layer (i.e. becoming a 2.5 layer) but has SAPs for all the intended areas of interaction. In this work's case it can be supposed that the UDP application communicates with the MIH through a high layer SAP.

The MIHF of the PoS was implemented containing the Event and Command services.

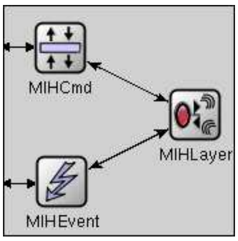

Figure B.3: MIH Layer for the PoS

The Event service in the PoS processes the Remote Events received from the MN. The information obtained through these events is then sent into the decision engine which, if necessary, issues commands through the Command service. These commands are related to handover initiation. Also, these are then sent to the UDP application, and then to the UDP layer where they are encapsulated in UDP packets and sent through the network to the MN.

B.3 Implementation Issues

This section describes some implementation issues that were encountered during the implementation phase of this thesis.

B.3.1 Authentication Failures in WLAN L2 Handover

As indicated in section B.1, WLAN L2 handover is comprised of a Authentication phase and an Association phase. In the first phase a problem was detected related to a slightly frequent chance of failing this authentication. The reason is not yet completely known, being only pointed to a timer in the MN going out before receiving the authentication response. In the normal model, prior to the required modifications to implement the 802.21 mechanism, the MN solved this issue by resending the authentication frame. But, in our case, to return to that state would involve resending all the signalling to the PoS. So, it was decided to implement a retry mechanism in the Wireless interface of the MN, where it retried the authentication three times. Upon failure, a new active scanning was conducted.

During the course of all the implementing and simulation, never a failure on association was reported and so, a retry mechanism for this phase was not necessary.

B.3.2 MIH Message selection for Handovers

A very important part of this project was the selection of which MIH messages should be implemented for NIHO. During the course of this project, several versions of the IEEE802.21 specification where analysed. Also, several signalling flow propositions where observed from the work-group members of 802.21 presentations. Nevertheless, even at this stage, the specification presents some inconsistency and lack of detail problems regarding some of the remote messages and their usage. For example, in the Annex section of the specification, a message flow chart exists but is inconsistent with the message descriptions in the previous chapters of the draft, and has several ambiguous behaviours. The main difficulties found in this process where:

- Lack of indication, in some messages, if they where local or remote;
- Lack of detail on the concrete effect of the reception of certain messages;
- Lack of indication of when certain messages should be sent;
- Ambiguous description of message function (for example, messages that where indicated as being remote, but their description was based on local behaviour.
- Message Parameter details missing.

The message flow contained in the Annex of the 802.21 specification was not helpful at all. In fact, the only message flow diagram that seemed to supply a proper set of MIH messages useful to NIHO was from an old 802.21 WG presentation, but that proposal had been reject.

A thorough analysis was done to each of the messages contained in the MIH message set, to select which ones where useful to our scenario. That selection represents our interpretation of the current specification draft.

B.3.3 Message Collision and Timers

A complicated factor that presented itself as a problem multiple times, was the occurrence of collision of packets involving MIH messages. This occurred mostly between MIH messages and video packets on the link path between the MN and the CN. At the time of starting the implementation, no reference was made on transport issues in the standard's draft. Without a mechanism that reacted to the loss, delay or collision of messages, the implemented MIH mechanism would be stalled, waiting for a message that would never arrive. Nevertheless, the IETF was aware of this issue in [41] and [37]. In the last document, two approaches where possible:

- Deal with transport issues in the MIH protocol;
- Rely on the underlying network mechanisms to solve these issues.

The first approach, if used, would end up recreating lots of already existing mechanisms in network functioning. The second, had the problem of being not as simple as if the issues where solved by the first item. Nevertheless, even if a simple retry mechanism for retransmission would be implemented, there where no concrete values on which should be the timers, how would other parameters such as speed of the terminal and RTT of the links affect them, etc.

To solve the collision issue two measures where adopted. First, the simulation network topology was simplified in order to reduce to the maximum the effort on links where MIH and video packets where passing. Secondly, the delay in these links was decreased, both to reflect the simulation of a viable cable length, and to reduce the delay of packets reaching the end points. The delay was configured to be 0.4ms.

B.3.4 Link Parameter Report Repetition

In order for operational mode B to work, several alterations where required in the already implemented framework. Since the signalling would be triggered at -80dBm, the MN would need to be able to move inside of the cell without sending any more Link Parameter Reports. This happened because the code in the Wireless Ethernet Module in OMNeT++, regarding analysis of the signal strength, is a cycle that always repeats itself. The condition for sending a Link Parameters Report was intro ducted there, so it was necessary to have a stop condition to prevent from sending this message at the end of each cycle. An indication that the message was already sent was needed to be implemented in the Wireless Ethernet Network Interface code, which prevented the resending of the message. As a consequence of this measure, it became necessary to reset this indication when the need for sending a proper Link Parameter Report presented itself, i.e. when entering a new cell or exiting and then re-entering the same cell. In the first case, the MN analyses the MAC addresses of the probe responses, while doing Active Scanning. When a probe response is received, the MAC address is stored. When a probe response with a different MAC is received, it means the MN entered a different cell, and so, the sending of a Link Parameter Report can be allowed. In respect to the second case, it required a more precise indication, since the terminal is re-entering the same cell and the MAC address is the same. In this case, a flag was introduced that becomes active whenever the MN finds an AP through Active Scanning. While doing active scanning, if the signal level was verified to be below the -80dBm threshold, and the flag was active, it meant the MN was leaving the cell, and so, the indication for sending Link Parameters Report is reset.

B.3.5 Link Detected Repetition

The same problem introduced with the Link Detected message. But with the introduction of the out-of-cell evaluation mentioned in the previous section, it became possible to only send one Link Detected per cell entry.

B.3.6 Out of Cell mechanism for MN localisation

This mechanism was implemented to decrease the number of failed handover initiations sent by the PoS, because when the message arrived at the MN, it was already unassociated or associated in a different AP. The mechanism consists in sending a Link Parameters Report message to the PoS, indicating that the signal level of the current AP has crossed the Association Threshold. This allows the PoS to know that the MN has left the current cell and it will not send unsolicited handover initiation commands, until the MN re entered the cell of an AP with acceptable load value.

B.3.7 UDP Applications for the PoS and for the MN

In order to build the MIH messages used in the signalling exchange for the handover preparation, two separate UDP applications had to be implemented, one for the MN and the other for the PoS. These UDP applications received indications and information from the respective MIH Functions, and created the remote MIH Events and Commands, and forwarded them to the UDP layer. Also, when the node was receiving MIH messages, it had to preserve the parameters contained in the MIH messages, and supply them to the MIH Function.

B.3.8 UDP Layer modification

In OMNeT++, it was discovered that there was no way to distinguish received packets from one another. So, a sequence number mechanism was implemented in the UDP Layer. This mechanism was later improved to be able to distinguish sequence numbers from packets containing MIH Messages, and from packets containing the video stream. This was necessary because packet loss was studied and so it was necessary to account for lost packets.

B.3.9 Simulation Delay and Parameter Configuration

One of the greatest hurdles imposed on simulation is the simulation time. Depending on the set of configurable parameters, a simulation can take quite some time. Also, if simulations with variable parameters are necessary, it is not feasible to manually configure all the parameters between simulation runs, since the idle time would be enormous, not taking advantage of the processing power.

OMNeT++ contains a mechanism to run several simulation runs in one machine, changing simulation parameters. Nevertheless, these parameters have to be configurable from within the omnetpp.ini file. If, for example, a more static parameter needs to be configured, this approach does not work. For example, the terminal speed is not configurable in the omnetpp.ini file.

Also, only using one computer is not feasible to do all the processing required. Time is a valuable asset regarding simulation work, and the more combined processing power, the better. OMNeT++ supports experimental distributed simulation, but it is too much complicated requiring special coding during implementation.

Lastly, OMNeT++'s statistical tools are not powerful enough to do all the required calculations like, for example, confidence intervals.

Solutions to automate simulations and to process results accordingly to the necessary work had to be devised.

B.3.10 Load Generation

The load generation for each of the APs was built using a structure maintained in the PoS. A birth and death Poisson process was implemented that created "virtual" nodes that arrived to the APs, moved around the cell, and then leave. A cap for the maximum number of allowed "virtual" nodes (as well as the MN) was also implemented. This mechanism simulated the existence of a MIH Function residing in the access points, in which an event would indicate to the PoS the number of nodes inside of the network.

B.3.11 Command-line Simulation and BASH scripting

To speed up simulations, the implemented code was compiled with no graphical interface. BASH scripts where built that parsed the configuration files between simulations, and changed them according to objectives set for this simulation. This was rather useful since it was required to simulate different speeds, thresholds, RTTs, delays, duration, random seeds, etc.

B.3.12 Result processing with BASH and OCTAVE scripting

The output of OMNeT++ simulations can be obtained trough vector and scalar files. OMNeT++ statistical tools where not powerful enough to process the data accordingly to the amount of required detail. Also, the command-line output generated in the simulations can provide valuable end-of-simulation results from calculations computed in the end of the simulation. For example, average time or number of handovers. Manual selection of the values to study would be impossible, taking into consideration the amount of collected data per simulation. It was resorted to the use of OCTAVE, which is a mathematical suite which can be run using command-line expressions, posing a powerful combination with BASH scripts. These outputs where parsed and, depending on the simulation study target, where feed to OCTAVE scripts responsible to calculate means and confidence intervals.To speed up this data processing, another script was made to remove the simulation results where no handovers where produced. Lastly, all results where combined in a single result file, which could be visualised with a text tool, our graphical analyser.

B.3.13 Graphic Result Construction using MATLAB and EXCEL

MATLAB scripts where built to parse the result files and create graphs for visual presentation of the results. Also, tables where made in EXCEL for numerical analysis of the results.

Bibliography

[1] IEEE, “Draft IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services,” 2006.

[2] T. Melia, A. de la Oliva, I. Soto, D. Corujo, A. Vidal, and R. Aguiar, “Impact of heterogeneous network controlled handovers on multi-mode mobile device design,” in *EEE Wireless Communications and Networking (WCNC), Hong Kong*, March 2007.

[3] M. Almeida, D. Corujo, S. Sargento, and R. Aguiar, “An End-to-End QoS Framework for 4G Mobile Heterogeneous Environments,” in *OpenNET Workshop, Diegem/Brussels, Belgium*, March 2007.

[4] V. Jesus, M. Almeida, D. Corujo, N. Senica, S. Sargento, and R. Aguiar, “Mobility with Quality-of-Service Support for Multi-Interface Terminals: Combined User and Network Approach,” in *IEEE Symposium on Computers and Communications (ISCC’07)*, 2007.

[5] T. Melia, A. de la Oliva, I. Soto, D. Corujo, A. Vidal, and R. Aguiar, “Toward ip converged heterogeneous mobility: A network controlled approach,” in *Under review in Computer Networks Journal*, December 2006.

[6] D. Corujo, A. Matos, T. Melia, J. Abeille, and R. Aguiar, “Problem Statement for Common Interface Support in Localized Mobility Management,” Internet Draft (Work in Progress), February 2007. [Online]. Available: http://www.ietf.org/internet-drafts/draft-corujo-ps-common-interfaces-lmm-00.txt

[7] J. Zander, “Radio resource management in future wireless networks - requirements and limitations,” in *IEEE Communications Magazine, 5(8), August 1997*, 1997. [Online]. Available: citeseer.ist.psu.edu/zander97radio.html

[8] A. Tolli, P. Hakalin, and H.Holma, “Performance evaluation of common radio resource management (CRRM),” in *International Conference on Communications, vol.5, no.pp. 3429- 3433*. IEEE, 2002.

[9] P. Magnusson, J. Lundsjo, J. Sachs, and P. Wallentin, “Radio resource management distribution in a beyond 3G multi-radio access architecture,” in *Global Telecommunications Conference, vol.6, no.pp. 3472- 3477 Vol.6*. IEEE, 2004.

[10] IEEE, “IEEE Standard for Local and Metropolitan Area Networks, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” 1999.

[11] ——, “EEE Standard for Local and Metropolitan Area Networks: Part 16 Air Interface for Fixed Broadband Wireless Access Systems,” 2004.

[12] ——, “IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1,” 2006.

[13] 3rd Generation Partnership Project, “Technical Specification Group Services and Systems Aspects; Network architecture,” 3GPP, March 2006.

[14] ——, “Technical Specification Group Core Network and Terminals; Handover procedures,” 3GPP, March 2006.

[15] Technical Specification Group Services and System Aspects, “3GPP System Architecture Evolution: Report on Technical Options and Conclusions,” 2006.

[16] 3rd Generation Partnership Project, “3GPP system to Wireles Local Area Network (WLAN) interworking; System description,” 3GPP, December 2006.

[17] A. Dutta, Y. Ohba, H. Yokota, and H. Schlzrinne, “Problem Statement for Heterogeneous Handover,” Internet Draft (Work in Progress), February 2006. [Online]. Available: http://www.ietf.org/internet-drafts/draft-ohba-mobopts-heterogeneous-requirement-01.txt

[18] T. Melia, J. Korhonen, and R. Aguiar, “Network initiated handovers problem statement,” Internet Draft (Work in Progress), March 2006. [Online]. Available: http://www.ietf.org/internet-drafts/draft-melia-mobopts-niho-ps-00.txt

[19] T. Melia, A. de la Oliva, I. Soto, P. Serrano, and R. Aguiar, “Network controlled handovers: challenges and possibilities,” in *Wireless Personal Communications Journal*, January 2007.

[20] D. Johnson, C. Perkins, J. Arkko, and T. Henderson, “Mobility Support for IPv6,” in *RFC 3775*. IETF, 2004.

[21] R. Koodli, “Fast Handovers for Mobile IPv6,” in *RFC 4068*. IETF, 2005.

[22] H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier, “Hierarchical Mobile IPv6 Mobility Management (HMIPv6),” in *RFC 4140*. IETF, 2005.

[23] P. Bertin, K. Guillouard, and J.-C.Rault, “IP based network controlled handover management in WLAN access networks,” in *International Conference on Communications, vol.7, no.pp. 3906-3910*. IEEE, 2004.

[24] K. Nishida, S. Isobe, T. Yagyu, and I. Akiyoshi, “Implementation and evaluation of a network-controlled mobility management protocol (IP2MM): performance evaluation compared with Mobile IPv6,” in *Wireless Communications and Networking Conference, vol.3, no.pp. 1402-1408*. IEEE, 2005.

[25] R. Aguiar and et al., “Scalable QoS-Aware Mobility for Future Mobile Operators,” in *IEEE Communications Magazine, Special issue on scaling the Mobile Internet*, June 2006.

[26] R. Moskowitz, P. Nikander, and P. Jokela, “Host Identity Protocol,” in *draft-ietf-hip-base-05, Internet Draft, work in progress*. IETF, 2006.

[27] A. G. Valko, “Cellular ip - a new approach to internet host mobility,” in *ACM Computer Communication Review*. ACM, January 1999.

[28] J. Kempf, K. Leung, P. Roberts, K. Nishida, G. Giaretta, and M. Liebsch, "Problem Statement for IP Local Mobility," in *draft-ietf-netlmm-nohost-ps-05, Internet draft, work in progress*. IETF, 2006.

[29] ——, "Goals for Network-based Localized Mobility Management," in *draft-ietf-netlmm-nohost-req-05, Internet draft, work in progress*. IETF, 2006.

[30] D. Farinacci, T. Li, S. Hanks, D. Meyer, and P. Traina, "Generic Routing Encapsulation," in *RFC 2784*. IETF, March 2000.

[31] E. C. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture," in *RFC 3031*. IETF, January 2001.

[32] S. Gundavelli, "Proxy Mobile IPv6," Internet Draft (Work in Progress), March 2007. [Online]. Available: http://www.ietf.org/internet-drafts/draft-sgundave-mip6-proxymip6-02.txt

[33] A. Bedekar, "A Protocol for Network-based Localized Mobility Management," Internet Draft (Work in Progress), March 2007. [Online]. Available: http://www.ietf.org/internet-drafts/draft-singh-netlmm-protocol-02.txt

[34] R. L. Aguiar and S. S. et all, "Scalable QoS-Aware Mobility for Future Mobile Operators," in *IEEE Communications Magazine*, 2006.

[35] M. Liebsch, A. Singh, H. Chaskar, D. Funato, and D. Funato, "Candidate Access Router Discovery (CARD)," in *RFC 4066*. IETF, 2005.

[36] J. Abeillé, R. Aguiar, T. Melia, I. Soto, and P. Stupar, "MobiSplit: a scalable approach to emerging mobility networks," in *MobiArch, San Francisco, USA*, December 2006.

[37] S. Sreemanthula, S. Faccin, E. Hepworth, and G. Daley, "Problem Statement and Requirements for Event and Command Services in Media Independent Handovers," in *draft-sreemanthula-es-cs-problem-02, Internet draft, work in progress*. IETF, 2006.

[38] S. Faccin et al, "Some Requirements for a Handover Information Service," in *draft-faccin-mih-infoserv-01, Internet draft, work in progress*. IETF, 2006.

[39] S. Park et al, "DHCPv4 Option for Discovering IEEE 802.21 Information Service Location," in *draft-daniel-dhc-mihis-opt-02, Internet draft, work in progress*. IETF, 2006.

[40] E. Hepworth et al, "Design Considerations for the Common MIH Protocol Functions," in *draft-hepworth-mipshop-mih-design-considerations-01, Internet draft, work in progress*. IETF, 2006.

[41] A. Rahman et al, "Transport of Media Independent Handover Messages Over IP," in *draft-rahman-mipshop-mih-transport-01, Internet draft, work in progress*. IETF, 2006.

[42] T. Melia et al, "Mobility Independent Services: Problem Statement," in *draft-melia-mipshop-mobility-services-ps-01, Internet draft, work in progress*. IETF, 2006.

[43] S. Zvanovec, M. Valek, and P. Pechac, "Results of indoor propagation measurement campaign for WLAN systems operating in 2.4 GHz ISM band," in *Antennas and Propagation, 2003. (ICAP 2003). Twelfth International Conference on*, 2003.

[44] M. Lott and I. Forkel, "A multi-wall-and-floor model for indoor radio propagation," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 1, May 2001, pp. 464–468.

[45] R. Chakravorty, P. Vidales, K. Subramanian, I. Pratt, and J. Crowcroft, "Performance Issues with Vertical Handovers-Experiences from GPRS Cellular and WLAN Hot-spots Integration," in *Pervasive Computing and Communications, 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on*. IEEE, 2004.

[46] R. Chakravorty and I. Pratt, "Performance Issues with General Packet Radio Service," in *Journal of Communications and Networks (JCN)*, 2002.

[47] P. Vidales, C. J. Bernardos, G. Mapp, F. Stajano, and J. Crowcroft, "A Practical Approach for 4G Systems: Deployment of Overlay Networks," in *First International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, 2005. Tridentcom 2005*, Trento, ITALY, February 2005, pp. 172 – 181.

[48] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli, "Integration of 802.11 and Third-Generation Wireless Data Networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*. IEEE, 2003.

[49] A. Dutta, S. Das, D. Famolari, Y. Ohba, K. Taniuchi, T. Kodama, and H. Schulzrinne, "Seamless Handoff across Heterogeneous Networks - An 802.21 Centric Approach," in *IEEE WPMC*, 2005.

[50] E. Wu, J. Lai, and A. Sekercioglu, "An Accurate Simulation Model for Mobile IPv6 Protocol," in *Proceedings of Australian Telecommunications, Networks and Applications Conference ATNAC 04*, December 2004.

[51] C. J. Bernardos, I. Soto, J. Moreno, T. Melia, M. Liebsch, and R. Schmitz, "Experimental evaluation of a handover optimization solution for multimedia applications in a mobile IPv6 network," *European Transactions on Telecommunications*, vol. 16, no. 4, pp. 317–328, April 2005.

[52] J. Lei, R. Yates, L. Greenstein, and H. Liu, "Wireless Link SNR Mapping Onto An Indoor Testbed," in *First International Conference on Testbeds and Research Infrastructures for the Development of Network and Communities (TRIDENTCOM'05)*. IEEE Computer Society, 2005.

[53] "Universal Mobile Access (UMA) User Perspective (Stage 1) R 1.0.0."

[54] A. Vidal, T. Melia, and A. Banchs, "Support for Centrally Coordinated Network Initiated Handovers," in *Contribution to IEEE P802.21/D02.00, 21-06-0783-00-0000-centralized_-NIHO*, October 2006. [Online]. Available: {http://www.ieee802.org/21/doctree/2006-11_meeting_docs/}

[55] D. Lister, S. Dehghan, R. Owen, and P. Jones, "UMTS capacity and planning issues," in *First International Conference on 3G Mobile Communication Technologies*. IEEE, 2000.

[56] T. Melia, A. de la Oliva, A. Vidal, C. Jesus, and I. Soto, "Analysis of the effect of mobile terminal speed on WLAN/3G vertical handovers," in *Wireless Communication Simposium, Globecom, San Francisco, USA*, November 2006.

[57] A. de la Oliva, T. Melia, A. Vidal, C. Jesus, I. Soto, and A. Banchs, "A case study: IEEE 802.21 enabled mobile terminals for optimized WLAN/3G handovers," in *Mobile Computing and Communication Review*, 2007.

[58] "Universal Mobile Access (UMA) Architecture (Stage 2) R 1.0.4."

[59] J. Laganier, "Network-based localized mobility management interface between mobile node and access router," Internet Draft (Work in Progress), June 2006. [Online]. Available: http://www.ietf.org/internet-drafts/draft-singh-netlmm-mn-ar-if-01.txt

[60] R. Giaffreda, F. Carvalho, and T. Melia, "Network enablers for enhanced mobility management," in *BT Technology Journal*, December 2006.

[61] Y. Ohba, "A framework of media-independent pre-authentication (mpa)," Internet Draft (Work in Progress), March 2007. [Online]. Available: http://www.ietf.org/internet-drafts/draft-ohba-mobopts-mpa-framework-04.txt

[62] 3rd Generation Partnership Project, "Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS); Service description; Stage 2," 3GPP, June 2006.

[63] ——, "Technical Specification Group Services and System Aspects; Functional stage 2 description of Location Services (LCS)," 3GPP, June 2006.

[64] ——, "Technical Specification Group RAN; UTRAN functions, examples on signalling procedures," 3GPP, June 2006.

[65] ——, "Mobile Application Part (MAP) specification," 3GPP, December 2006.

[66] ——, "General Packet Radio Service (GPRS); GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface," 3GPP, December 2006.

[67] ——, "Serving GPRS Support Node SGSN - Visitors Location Register (VLR); Gs Interface Network Service Specification," 3GPP, July 2005.

[68] ——, "General Packet Radio Service (GPRS); Base Station System (BSS) - Serving GPRS Support Node (SGSN) interface; Gb interface Layer 1," 3GPP, February 2005.

[69] ——, "General Packet Radio Service (GPRS); Base Station System (BSS) - Serving GPRS Support Node (SGSN) interface; Network service," 3GPP, November 2006.

[70] A. Patel, K. Leung, M. Khalil, and H. Akhtar, "Authentication Protocol for Mobile IPv6," in *RFC 4285*. IETF, January 2006.

[71] A. Patel, K. Leung, M. Khalil, H. Akhtar, and K. Chowdhury, "Mobile Node Identifier Option for Mobile IPv6," in *RFC 4283*. IETF, November 2005.

[72] B. Patil, P. Roberts, and C. E. Perkins, "IP Mobility Support for IPv4," in *RFC 3344*. IETF, August 2002.

[73] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," in *RFC 2460*. IETF, 1998.

[74] T. Melia et al, “Mobility Independent Services: Problem Statement,” in *draft-ietf-mipshop-mis-ps-00, Internet draft, work in progress*. IETF, 2007.

[75] T. Narten, E. Nordmark, and W. Simpson, “Neighbor Discovery for IP Version 6 (IPv6),” in *RFC 2461*. IETF, 1998.

[76] A. Banchs et al., “End-to-end Delay Analysis and Admission Control in 802.11 DCF,” in *Computer Communications*. Elsevier, 2006.

[77] P. de Silva and H. Sirisena, “A mobility management protocol for ip-based cellular networks,” in *ICCN 2001*. to check, to check 2001, p. 000.

[78] A.-E. M. Taha, H. S. Hassanein, and H. T. Mouftah, “On robust allocation policies in wireless heterogeneous networks,” in *First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks*. IEEE, June 2004, pp. 198–205.

[79] “Technical specification group ran: Signalling enhacements for circuit-switched (cs) and packet-switched (ps) connections;analyses and recommendations(release 7).” 3GPP.

[80] The Daidalos Project, “Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimised personal Services,” in *http://www.ist-daidalos.org/*, 2006.

[81] R. Aguiar, D. Bijwaard, B.Farshichian, K. Jonas, and A. Sarma, “Pervasive services for next generation heterogeneous networks,” in *WTC06, World Telecommunications Congress 2006*, May 2006.

[82] T. Melia, R. Aguiar, P. Pacyna, and A. Banchs, “Niho: Network initiated handovers for next generation all ip networks,” in *Demo/Poster Session, Infocom, Barcelona, Spain*, April 2006.

[83] M.Williams, “Directions in Media Independent Handover,” in *IEICE Transactions on Fundamentals of Electronics*. Communications and Computer Sciences, 2005.

[84] S. Sargento, V. Jesus, M. Almeida, D. Corujo, F. Sousa, F. Mitrano, T. Strauf, J. Gozdecki, and G. Lemos, “Context-Aware End-to-End QoS Architecture in Multi-technology, Multi-interface Environments,” in *16th Mobile and Wireless Communications Summit (MobileSummit 2007)*, 2007.