DEPARTMENT OF INFORMATION &
COMMUNICATION SYSTEMS
ENGINEERING

COMPUTER & COMMUNICATION
SYSTEMS LABORATORY

MSc THESIS

COMMUNICATION PROTOCOLS FOR SEAMLESS SERVICE PROVISIONING IN COEXISTING HETEROGENEOUS WIRELESS TECHNOLOGIES

Written by:
PRODROMOS MAKRIS

SUPERVISOR:
CHARALABOS SKIANIS, ASSISTANT PROFESSOR

KARLOVASSI SAMOS, GREECE, FEBRUARY 2009

EXECUTIVE SUMMARY

It has been a constant human endeavour to communicate more effectively and at the same time to be free of any bondage, physical or psychological. A similar underlying trend can also be seen in the evolution of telecommunications. The need for mobility and higher speeds in an ever-changing environment has been of paramount importance. The next generation of mobile and wireless systems will more than likely converge and integrate architectural, network and application levels. The unification of 3G, WLANs, WPANs, WiMAX, satellite systems as well as other emerging access technologies, will enable users to access common services using TCP/IP running over them.

Many researches have been studied for the coexistence of different wireless technologies as a general network model in the next-generation networks. The research trends can be divided into two groups: one trend focuses on the integration of distinct networks, while the other trend focuses on the supportability of mobility in interworking networks. The former research trend has been studied for a long time, and there are many research results with three main integration types, tight, loose and hybrid integration. On the contrary, the latter trend has only recently began to be studied. Especially, the mobility issue with the 802.21 framework has been considered as a reasonable solution in the next-generation networks, since the integration of networks such as UMTS/802.16e/802.11 demands a robust, reliable and efficient framework. IEEE P1900.4 is also an emerging standard for optimized radio resource utilization where cognitive radio technologies are used for efficient spectrum utilization. These two standards, which will be thoroughly investigated in this thesis, provide communication protocols in order to accomplish seamless services for the end user.

In this work, we try to present a thorough survey of two supplementary draft standards, IEEE 802.21 and IEEE P1900.4, which promise to solve plenty of the open issues in the seamless service provisioning research area in coexisting heterogeneous wireless technologies. These two draft standards are separately developed by two different working groups. We consider that with a specific cooperating architecture, we can exploit the architectural building blocks of both standards. Therefore, we will be able to accomplish optimum vertical handover procedures (802.21) as well as making optimum radio resource usage of the available spectrum (P1900.4). Thus, we present real use case scenarios, accompanied by specified flow of messages among the architectural building blocks in order to prove our assertions.

CONTENTS

CHAPTER 1 - 4G ERA IN A NUTSHELL

CHAPTER 2 - IEEE 802.21 STANDARD FOR MEDIA INDEPENDENT HANDOVER SERVICES

CHAPTER 3 - USE CASES FOR HANDOVER PROCEDURES BASED ON 802.21

CHAPTER 4 - IEEE P1900.4: RECONFIGURATION OF MULTI-RADIO SYSTEMS

CHAPTER 5 - REAL USE CASE SCENARIOS BASED ON P1900.4 DRAFT STANDARD

CHAPTER 6 - 802.21 & P1900.4 STANDARDS COEXISTING IN A FUTURE NETWORK ARCHITECTURE

LIST OF FIGURES

CHAPTER 1
4G ERA IN A NUTSHELL

1.1 WIRELESS NETWORK TECHNOLOGIES COMPETING IN THE NEXT GENERATION ERA

Many researches have been studied for the coexistence of different wireless technologies as a general network model in the next-generation networks. The research trends can be divided into two groups: One trend focuses on the integration of distinct networks, while the other trend focuses on the supportability of mobility in interworking networks [1]. The former research trend has been studied for a long time, and there are many research results with three main integration types, tight, loose and hybrid integration. A Multi-Scenario Based Call Admission Control for Coexisting Heterogeneous Wireless Technologies is examined in [2, 3]. The reader can find plenty of information in these two works. On the contrary, the latter trend has only recently began to be studied. Especially, the mobility issue with the 802.21 framework has been considered as a reasonable solution in the next-generation networks, since the integration of networks such as UMTS/802.16e/802.11 demands a robust, reliable and efficient framework. IEEE P1900.4 is also an emerging standard for optimized radio resource utilization where cognitive radio technologies are used for efficient spectrum utilization. These two standards, which will be thoroughly investigated in this thesis, provide communication protocols in order to accomplish seamless services for the end user. In this chapter, we try to present in a sententious manner all the wireless network technologies that are going to compete in the 4G market. More specifically, UMTS, HSPA, WLAN, WiMAX, WPAN and DVB infrastructures are going to be demonstrated.

1.1.1 UMTS – The 3G network

Often termed the holy grail of wireless capabilities, advanced mobile wireless, or third-generation services, 3G systems allow for high-speed, always-on data transmission. They are intended to provide access to a wide range of telecommunications services, specifically for mobile users. Worldwide roaming and services capability, Internet, and other multimedia applications are some of the key features that fit the 3G paradigm. For 3G data applications, the specifications generally require the ability to support both circuits and packets with the following rates and caveats: a) 144 Kbps (minimum) or higher for vehicular traffic, b) 384 Kbps (minimum) for pedestrian mobile users, c) 2 Mbps or more for indoor, semi-stationary mobile users, d) asymmetric data rates for send and receive, e) multimedia calls including video calls, f) both fixed and variable rate data traffic [4].

The UMTS network standard is recognized as a third-generation (3G) function set. It addresses the growing demand for new mobile and Internet applications and for added capacity in the overcrowded mobile communication sky. It uses a pair of 5

MHz wide radio frequency channels, selecting the uplink from the 1900 MHz band and the downlink from the 2100 MHz band. UMTS is envisioned to allow many more applications to be introduced to a worldwide base of users and provides a vital link between existing multiple GSM/GPRS systems. Simply put, it is the ultimate single worldwide standard for all mobile telecommunications. Thus, referring to cellular systems in the 4G era, we substantially mean UMTS standard and vice versa.

For more details about WCDMA, the architecture and various QoS issues about UMTS, the reader can see through our undergraduate thesis (chapter 2) in [2].

1.1.2 HSPA – The 3.5G network

HSPA is the successor of UMTS in the mobile telecommunications field. We separate it from the basic 3G standard, as it seems to become very popular in the near future because of its “4G-like” characteristics. Some more optimistic considerations refer that HSPA will become the master of the future cellular market. In this subsection we refer in a sententious manner some basic characteristics of HSPA, which are supplementary to those we already know from UMTS standard.

HSPA implements a fast and complex channel control mechanism with new adaptive modulation and coding techniques and a fast scheduler. HSPA works within a WCDMA downlink 5 MHz channel and is targeted at 0.9 to 10 Mbps data rates with the current standard allowing up to 14.4 Mbps. HSDPA (the downlink part of HSPA) is targeting 20 Mbps in release 6 of the 3GPP standardization process using an option called MIMO technique. Nowadays, HSDPA deployments target up to 10 Mbps per user. These data rates compete wireline DSL and cable modem technology. This places HSDPA data capabilities into the 3.5G range [4].

High-speed downlink packet access (HSDPA) was standardized as part of 3GPP Release 5 with the first specification version in March 2002. High-speed uplink packet access (HSUPA) was part of 3GPP Release 6 with the first specification version in December 2004. HSDPA and HSUPA together are called high-speed packet access (HSPA). The first commercial HSDPA networks were available at the end of 2005 and the commercial HSUPA networks are already available since 2007. The HSDPA peak data rate available in the terminals was initially 1.8 Mbps and increased to 3.6 and 7.2 Mbps during 2006 and 2007, and potentially beyond 10 Mbps in 2008 and in the near future. The HSUPA peak data rate in the initial phase was 1–2 Mbps with the second phase pushing the data rate to 3–4 Mbps. In figure 1.1, radio capability evolution from GPRS to HSPA is illustrated.

HSPA is deployed on top of the WCDMA network either on the same carrier or (for a high-capacity and high bit rate solution) using another carrier. In both cases, HSPA and WCDMA can share all the network elements in the core network and in the radio network including base stations, Radio Network Controller (RNC), Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). WCDMA and HSPA are also sharing the base station sites, antennas and antenna lines. The upgrade from WCDMA to HSPA requires new software package and, potentially, some new pieces of hardware in the base station and in RNC to support the higher data rates and capacity. Because of the shared infrastructure between WCDMA and HSPA, the cost of upgrading from WCDMA to HSPA is very low compared with building a new standalone data network [5].

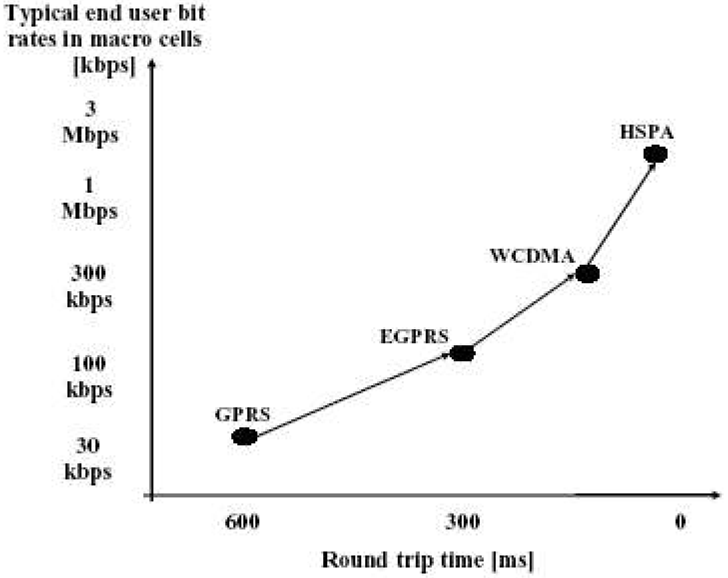


Figure 1.1 : Radio capability evolution in cellular networks [5]

The first HSDPA terminals are data cards providing fast connectivity for laptops. It is expected that HSPA will be a standard feature of most 3G terminals after some years in the same way as Enhanced Data Rates for GSM Evolution (EDGE) capability is included in most GSM/GPRS terminals. HSDPA will also be integrated to laptop computers in the future, as is indicated already by some of the laptop manufacturers.

Higher cell capacity and higher spectral efficiency are required to provide higher data rates and new services with the current base station sites. Basic HSPA includes a one-antenna Rake receiver in the terminals and two-branch antenna diversity in the base stations. Enhanced HSPA includes two-antenna equalizer mobiles and interference cancellation in the base station. The simulation results show that HSPA can provide substantial capacity benefit. Basic HSDPA offers up to three times WCDMA downlink capacity, and enhanced HSDPA up to six times WCDMA. The spectral efficiency of enhanced HSDPA is close to 1 bit/s/Hz/cell. The uplink capacity improvement with HSUPA is estimated between 30% and 70%. HSPA capacity is naturally suited for supporting not only symmetric services but also asymmetric ones with higher data rates and volumes in downlink.

1.1.3 IEEE 802.11x - WLAN

Wireless LANs are creating attractive new growth opportunities. Simple to deploy and relatively inexpensive to acquire, WLAN technology uses unlicensed spectrum to achieve data transfer rates at 1 Mbps, 2 Mbps, 5.5 Mbps, 11 Mbps, 54 Mbps, and beyond. These data rates are significantly higher than the specifications of 3G and are potentially disruptive to 3G data service offering especially for pedestrian-based wireless services.

The early iterations of WLAN technology left much room for improvement with primary issues related to security and limited distance. Deployed in the ISM band of the radio spectrum at 2.4 GHz, there are caps and limitations on power and range to reduce interference with the surplus of other devices in this unlicensed area.

The 802.11x standard represents the technical specifications of WLANs. The 802.11b, 802.11a, and 802.11g standards are currently the most important. Given the statutory and physics-related constraints on the improvement vector of WLANs, or more specifically, the 802.11x standard, 3G data technologies and 802.11x technologies could end up as complementary products that can be knitted together as a blanket of coverage. A parallel for this is one where 802.11x would represent wireless LANs while 3G data technologies would characterize wireless WANs [4].

WLANs are more commonly referred to using the marketing term Wi-Fi, which is short for Wireless Fidelity. Using a Wi-Fi card or Wi-Fi integrated technology in laptops or handhelds, mobile computing users can surf the Internet at 11 Mbps up to 54 Mbps speeds without physically plugging their computer into anything, in this case, as long as they are within about 100 metres of a Wi-Fi hotspot's central access point (AP). Essentially, Ethernet through the air, Wi-Fi is a technology that is easily co-opted and overlaid on existing wired LANs to provide laptop computers more ubiquitous access and always-connected capability. Wi-Fi is cellular in concept, the Wi-Fi AP is generally at the center of a wireless cell that radiates the radio frequencies for a couple hundred metres in circumference. Deploying several APs and tying them together through a network backbone infrastructure enables wireless roaming from Wi-Fi cell to Wi-Fi cell.

The reader can find more details about 802.11 architecture, physical-layer techniques and QoS issues in our undergraduate thesis (chapter 3) in [2]. Moreover, in chapters 6 and 7 of the same work, the reader can find a survey-like work on the existing integration architectures in a UMTS/WLAN scenario. This will help the reader in understanding the general concept of our research goals and what was the trigger for this postgraduate thesis.

1.1.4 IEEE 802.16x - WiMAX

The prospect of broadband Internet access anywhere, at any time, has seemed a distant dream, far from reality for the vast majority of PC, laptop and handheld users. However, with WiMAX, it will soon become something users cannot live without. WiMAX is one of the hottest wireless technologies around today.

WiMAX systems are expected to deliver broadband access services to residential and enterprise customers in an economical way. Although it has one name, WiMAX will be two different market technologies. The first is for fixed wireless and falls under the IEEE 802.16-2004 standard approved in 2005. The second, for mobile applications, is under the 802.16e specification and has been approved in 2007. As of now, fixed WiMAX is capable of becoming a replacement for DSL or cable or for network backhaul. In future, WiMAX will transform the world of mobile broadband by enabling the cost-effective deployment of metropolitan area networks based on the

IEEE 802.16e standard to support notebook PC and mobile users on move. We will particularly focus on 802.16e, as we are interested in wireless access technologies.

There are many advantages of systems based on 802.16e, for example the ability to provide service even in areas that are difficult for wired infrastructure to reach and the ability to overcome the physical limitations of traditional wired infrastructure. The standard will offer wireless connectivity of up to 30 miles. The major capabilities of the standard are its widespread reach, which can be used to set up a metropolitan area network, and its data capacity of 75 Mbps. This high-speed wireless broadband technology promises to open new, economically viable market opportunities for operators, wireless Internet service providers and equipment manufacturers. The flexibility of wireless technology, combined with high throughput, scalability and long-range features of the IEEE 802.16e standard helps to fill the broadband coverage gaps and reach millions of new residential and business customers worldwide (Figure 1.2) [6].

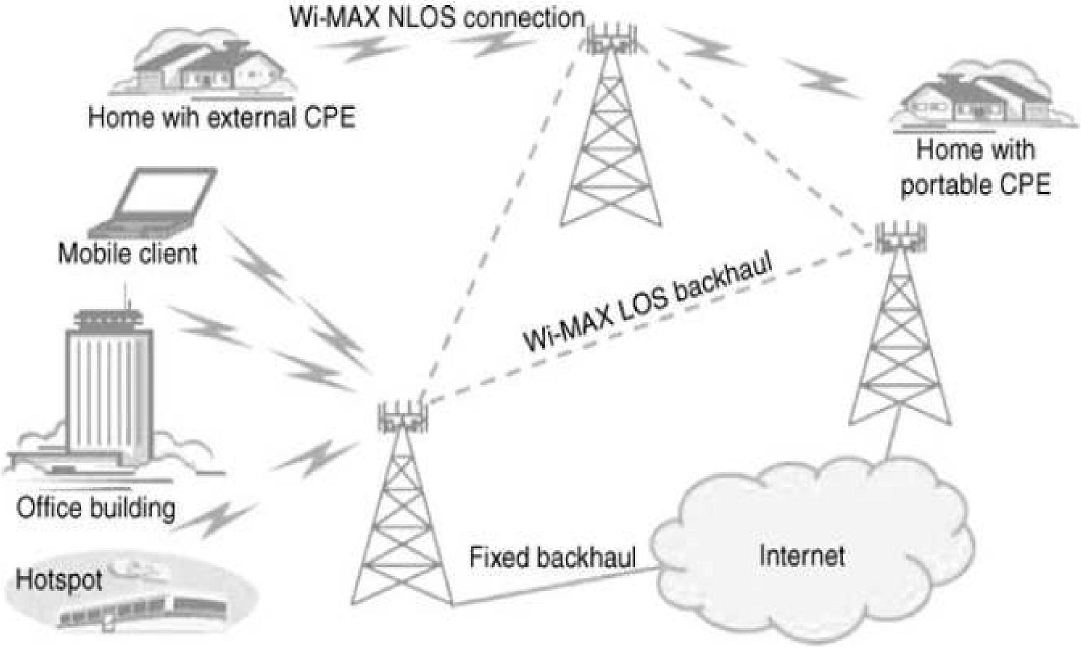


Figure 1.2 : WiMAX solutions [6]

WiMAX in a few words is: a) a wireless technology optimized for the delivery of IP centric services over a wide area, b) a certification that denotes interoperability of equipment built to the IEEE 802.16 or compatible standard, c) a scalable wireless platform for constructing alternative and complementary broadband networks.

1.1.5 IEEE 802.15x - WPAN

A WPAN is a network solution that enhances our personal environment, either work or private, by networking a variety of personal and wearable devices within the space surrounding a person, and providing the communication capabilities within that space and with the outside world. WPAN represents the person-centered network concept, which allows the person to communicate with his/her personal devices close to him/her (e.g., personal digital assistants, webpads, organizers, handheld computers, cameras, and head-mounted displays) and to establish the wireless connections with the outside world. The PAN searches for technical challenges and applications that will turn it into a system really serving the person and improving the quality of his/her personal and professional life. Many different operating scenarios can be foreseen, mainly concentrated around: a) personal services, b) business services and c) entertainment [7].

The personal services include medical telemonitoring and control applications as well as smart homes. Medical professionals can constantly monitor the sick person wherever he/she moves. Within his/her home a person can move, addressing different devices and getting useful information from them.

The business scenario includes building environmental monitoring, fleet management, and person searching. The content and temperature pressure in the truck could be measured, as well as the production line, the movement of the employees within the offices etc.

Entertainment scenarios include high-tech applications such as high-speed video on the trains, information retrieval through the personal devices, and other imaginable services. Applications will range from very low bit rate sensor and control data up to very high bit rate video streaming and bandwidth-demanding interactive games. The system should be able to support the variety and complexity of communication services and devices. The most capable devices should incorporate multimode and multi-homing functionalities that will enable the access to multiple networks. More details about these issues of crucial significance will be addressed in the next chapters of this work.

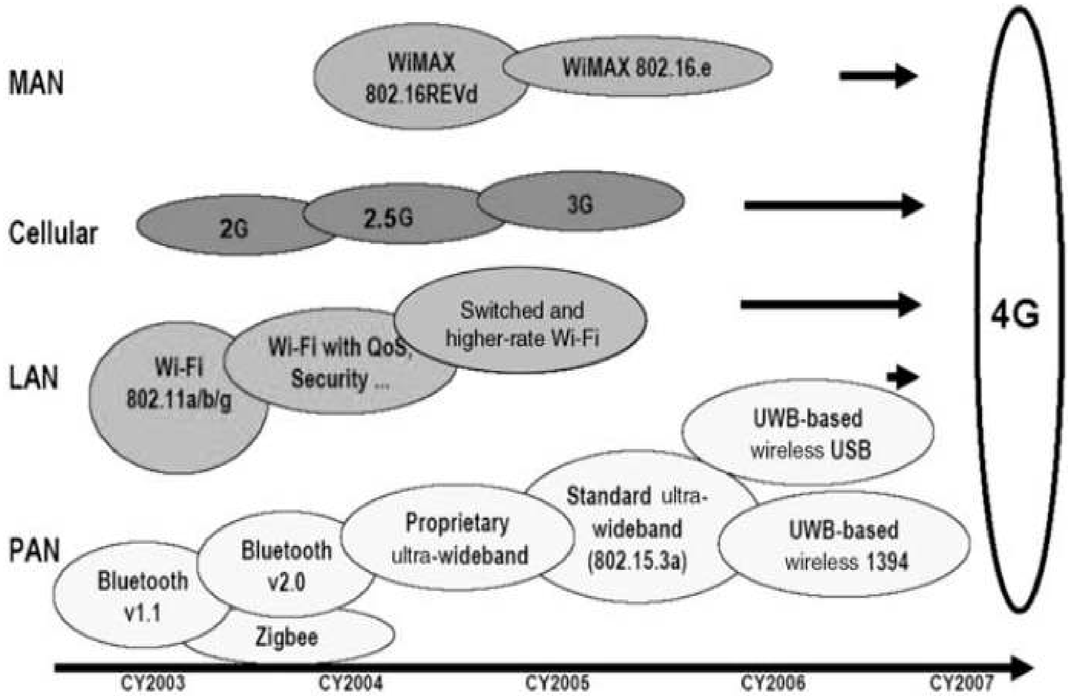


Figure 1.3: The evolution to 4G wireless network systems [6]

1.1.6 DVB – Satellite Systems

In satellite communications, a division is made between Fixed Satellite Service (FSS), Broadcast Satellite Service (BSS) and Mobile Satellite Service (MSS) delivery. In FSS, satellites operate mainly in a point-to-point mode as part of the core network and provide high-capacity links in telecommunications and ISPs. On a point-to-point basis, IPv6 operations pose no problems and are currently in operation over many satellite links. Within the FSS/BSS domain, satellites are used extensively to deliver video services to cable heads or Direct To Home (DTH). Interactive television is available in this digitized service through low rate channels with proprietary protocols via, mostly at the moment, landline. The DVB-S standard is becoming an industry standard for the delivery of IP via satellite, although it was not primarily designed for this purpose and is not optimal.

So far mobile satellites have either selected niche areas or tried to compete in the mass market with cellular services. In the long run, such competition will not be fruitful, but collaboration with cellular services in the access networks will be beneficial. This is true mainly in two areas. The first is in the coverage of remote areas that would be too expensive to be served by cellular services. Providing such services would be more expensive but could form a value-added offering for mobile service providers. An adjunct to this would be the provision of disaster services to back up cellular services. The second, and perhaps more interesting, is the delivery of Multimedia Broadcast and Multicast Service (MBMS) services to the mass market. Within 3G networks and also beyond, these services are very difficult and expensive to be delivered in a cellular format. However, they are ideally matched to the attributes of a satellite network in terms of the broadcast coverage and thus we have a win–win situation. The delivery of multimedia content in MBMS to a large customer base via satellite can reduce the cost by orders of magnitude over cellular networks. Moreover, with sufficient large storage capacity in the user terminals, unidirectional point-to-multipoint services are able to provide on-demand and interactive applications because push and store mechanisms make the point-to-multipoint relationship transparent to users.

In a truly integrated satellite/cellular system to be used by mobile operators, satellites will be complemented by terrestrial repeaters known as Intermediate Module Repeaters (IMRs) to provide cost-effective services. This collaboration is what is envisioned in a B3G environment. We will not deal with interworking architecture scenarios, which include DVB technologies in the present work. A little work has been done so far in the specific field by the international research community because of some inherent drawbacks of this wireless network technology. Both 802.21 and P1900.4 do not consider interworking scenarios including satellite systems, and thus we won’t deal with them in the rest of our work, leaving it as one of the topics of our future work.

1.2. THE ISSUE OF INTEGRATING HETEROGENEOUS WIRELESS SYSTEMS IN A 4G CONTEXT

There is no single technology that can provide the consumer with all the desired applications, but integration of mobile and wireless systems could complement each other. All these technologies being available, it is quite reasonable to consider that a fourth generation of mobile system will give us the ability to integrate these technologies even though they work on different platforms. As we have different techniques for communicating, an integrated system will have the ability to utilize all their advantages, so shortcomings of an individual technology will be complemented by other systems [6].

B3G systems will need to assimilate and integrate existing technologies, rather than supplant them. It is envisioned that present mobile cellular systems will be "blended into" B3G technology, which will enable mobile cellular devices to roam seamlessly from WMAN to WLAN, to WPAN and vice versa without difficulty.

A B3G network, a heterogeneously networked environment, will require that handheld devices evolve considerably, from the limited (often fixed function and fixed network) devices that predominate today, to powerful, flexible devices that can intelligently interact with multiple, heterogeneous networks and services. A universal communicator-class device is projected to be a flexible, powerful personal communication appliance that provides users with transparent access to any available network, at any time, including the ability to seamlessly roam across those networks. Such a device must also provide support for key usage models that are made possible by a mixed-network environment. These usage models include the following: a) info-fuelling (smart data transfers using best available/most appropriate network), b) simultaneous voice and data sessions, c) rich media that scales across networks (for example, video quality increases in a higher-bandwidth environment, d) cross-network voice, including support for seamless handoff and e) location-based services.

Enabling such ubiquitously connected devices poses numerous difficult technological challenges. These include the following [8]:
a) Multiple radio integration and coordination: the device integrates multiple radios.
b) Intelligent networking – seamless roaming and handoff: users can expect to roam within and between networks like they do with today's cell phones.
c) Power management: future handsets and other devices will run richer applications, and power management will become an even greater challenge.
d) Support for cross-network identity and authentication: providing a trusted and efficient means of establishing identity is one of the key issues in cross-network connectivity.
e) Support for rich media types: the addition of a high-bandwidth broadband wireless connection, such as a WLAN, will open up new opportunities for the delivery of rich media to handheld devices.
f) Flexible, powerful computing platform: the foundation of a universal communicator-class device must be a flexible, powerful, general-purpose processing platform.
g) Overall device usability: meeting all these challenges must not render the device "user-unfriendly".

Because the Internet and cellular systems were designed and implemented by people with different backgrounds in computers and communications, respectively, their integration will not be a simple task. Such integration, however, can be considered to be a first step toward B3G networks, where heterogeneous networks must work together in order to provide differentiated services to users in a seamless and transparent manner [8].

1.3. CHALLENGES IN THE MIGRATION TO 4G MOBILE SYSTEMS

The problem of small spectrum utilization is a major challenge to deal with in the next generation wireless systems. Fixed spectrum allocation proved to be inefficient. FCC measurements have indicated that 90% of the time, many licensed frequency bands remain unused. This gives opportunity for radio access networks and terminals to detect and use temporally unused spectrum, thus improving spectrum utilization. P1900.4 standard deals with these issues and it will be further investigated in chapters 4 and 5.

Another important problem is the excessive delay and jitter that is incurred by the often handoff procedures. In a 4G environment, we expect mobile terminals to move anywhere, anytime and thus performing several handoffs during their random movements. Therefore, vertical handoff from one system to another is a usual event. These handoff procedures inevitably incur excessive delays, causing remarkable problems at the application layer in the end user side. Delay sensitive services such as voice calls and video streaming suffer the most. Thus, we need specific communication protocols for exchanging information between network and terminal entities. IEEE 802.21 standard deals with media independent handover services and tries to solve the pre-referred challenges. It will be deeply investigated in chapters 2 and 3.

The introduction of multimedia services into mobile communications will require mobile transmission speeds of up to 100 Mbps. Therefore, a wider frequency band than that in 3G will have to be assigned to B3G mobile communication systems. Generally speaking, mobile communication systems should be assigned the lowest available frequency band when taking into account path loss in radio channels. However, it will probably be impossible to assign a lower frequency band than that of 3G to B3G systems because of the fact that these bands are already regulated and in use. Therefore, techniques that enable high-speed data transmission within a limited frequency band will become important in B3G systems. Simply put, techniques for increasing the efficiency of frequency utilization will play a great role in B3G systems.

In [9], there is a summary table of the major challenges that 4G has to overcome. These are distinguished in three categories, from the mobile station's, the system's and the service's perspective. In the first category, the major challenges are:
a) <u>multi-mode user terminals</u>: there is a need to design a single user terminal that can operate in different wireless networks and overcome the design problems such as

limitations in device size, cost, power consumption and backward compatibilities to systems,
b) wireless system discovery: need to discover available wireless systems by processing the signals sent from different wireless systems (with different access protocols and incompatible with each other),
c) wireless system selection: the proliferation of wireless technologies complicates the selection of the most suitable technology for a particular service at a particular time and place.

From the system's perspective, the major challenges are: a) terminal mobility issues, b) network infrastructure and QoS support, c) fault tolerance and survivability and d) security issues. Eventually, in the last category we need to face problems such as personal mobility and multi-operators and billing system issues.

1.4. MULTIMODE & RECONFIGURABLE PLATFORMS

A major contributor toward the convergence of platforms in the B3G era is reconfigurability, which provides technologies (SDRs) that enable terminals and network segments to dynamically adapt to the set of radio access technologies (RATs) that are most appropriate for the conditions encountered in specific service area regions and at specific times of the day. RAT selection is not restricted to those preinstalled in the elements. On the contrary, the missing components can be dynamically downloaded, installed, and validated. Reconfigurability poses requirements on the functionality of wireless networks. Some of the challenges that have to be met to realize the reconfigurability concept are given below in this section [8].

First, three families of scenarios that must be taken into account when designing the reconfigurability technology have been identified: the promises of ubiquitous access, pervasive services, and dynamic resources provisioning. Ubiquitous access is mainly targeted at increasing the worldwide access to services. It relates to the support of users who turn on a device in a wireless environment to which it has not been previously connected. Roaming is another example of this scenario, and the reconfigurability concept must increase roaming possibilities for users. The concept of pervasive services stresses the need for reconfigurability when several radio access technologies are present in a given wireless environment. Indeed, the proper use of these different access technologies and reconfigurable equipment needs many capabilities like system discovery, protocol reconfiguration, and a method of vertical handover. Dynamic resources provisioning involves a dynamic reconfiguration of the terminal and network elements to improve the bandwidth for users with better adapted radio interfaces as well as additional spectrum. In this case, the protocol stack must be updated in the terminal and in the network. Consequently, the different communication systems covering such areas must be able to adapt to load and services variations.

Second, reconfigurability research has identified the concept of a management and control system that enables network elements to operate in an end-to-end reconfigurability context. The main idea of this concept is a clear separation of network management and control functions. Reconfigurable components, like programmable processors and reconfigurable logic, are envisioned for reconfigurable equipment. B3G architecture needs to support the dynamic insertion and configuration of different protocol modules as devices join and leave the given wireless environment. Furthermore, the reconfigurability of SDR equipment is widely seen as one of the enabling technologies for communication systems beyond 3G.

Lastly, efficient spectrum management is of prime importance for reconfigurability to be realized. In discussions on reconfigurability, efficient spectrum management is one of the components of radio resource management, which also includes a joint management of radio resources belonging to different (2G and 3G) RATs with fixed spectrum allocation, cognitive radio, and a progressive network planning process. RRM is a complex process, but is necessary for the deployment of B3G networks. It consists of dynamically managing a spectrum as well as allocating traffic dynamically to the RATs participating in a heterogeneous, wireless access infrastructure. The coexistence of diverse technologies that form part of a heterogeneous infrastructure has brought about the idea of flexibly managing the spectrum. This implies that fixed frequency bands are no more guaranteed for RATs, but through an intelligent management mechanism, bands are allocated to RATs dynamically in a way that ensures that the capacity of each RAT is maximized and interference is minimized. Furthermore, there is a tight relationship between spectrum management and cognitive radio. Cognitive radio will provide the technical means for determining in real time the best band and the best frequency to provide the services desired by a user.

1.5 REFERENCES

[1] Joo-Young Baek et al, “**Network-initiated Handover Based on IEEE 802.21 Framework for QoS Service Continuity in UMTS/802.16e Networks**”, May 2008.

[2] Makris Prodromos, “**Thorough Investigation of Heterogeneous Wireless Network Technologies in the 4G era**”, Undergraduate thesis, Department of Information & Communication Systems Engineering, University of the Aegean, October 2007.

[3] Makris Prodromos & Charalabos Skianis, “**Multi-Scenario Based Call Admission Control for Coexisting Heterogeneous Wireless Technologies**”, Globecom 2008, New Orleans, 30 November – 4 December.

[4] Robert Wood, **Next Generation Network Services**, Cisco Press, 2006, Chapter 9.

[5] Harri Holma & Antti Toskala, “**HSDPA/HSUPA for UMTS**”, John Wiley & Sons Ltd., 2006, Chapter 1.

[6] Deepak Pareek, “**The Business of WiMAX**”, John Wiley & Sons Ltd., 2006, Chapter 1.

[7] Ramjee Prasad, Luis Munoz, “**WLANs and WPANs towards 4G Wireless”**, Artech House Inc., 2003, Chapter 6.

[8] Hsiao-Hwa Chen, Mohsen Guizani, “**Next Generation Wireless Systems and Networks**”, John Wiley & Sons Ltd., 2006, Chapter 6.

[9] http://www.comlab.hut.fi/opetus/333/2004_2005_slides/4G_text.pdf

CHAPTER 2
IEEE 802.21 STANDARD FOR MEDIA INDEPENDENT HANDOVER SERVICES

2.1 SCOPE – PURPOSE – OVERVIEW

The scope of the IEEE 802.21 standard is to facilitate the handover between IEEE 802 and non-IEEE 802 access networks (e.g. cellular networks) in a way that is independent from particular access network features (i.e. perform media-independent handover (MIH)). This is realized through its mechanisms that offer, for example, the ability to explicitly indicate an imminent break in communication or link deterioration. This report mechanism is implemented through specific triggers that convey useful information related to mobility to entities where a decision is made and in turn, cause a command to be executed at some specific network elements [1].

The main purpose of IEEE 802.21 is to enable handovers between heterogeneous technologies (including IEEE 802 and cellular technologies) without service interruption, hence improving the user experience of mobile terminals. Many functionalities required to provide session continuity depend on complex interactions that are specific to each particular technology. IEEE 802.21 provides a framework that allows higher levels to interact with lower layers to provide session continuity without dealing with the specifics of each technology. That is, the upcoming protocol can be seen as the "glue" between the IP-centric world developed in the IETF, and the reference scenarios for future mobile networks currently being designed in the 3GPP and 3GPP2 or other technology specific solutions. While the IETF does not address specific layer 2 technologies, the interest of 3GPP/3GPP2 in non-cellular layer 2 technologies, such as WLAN, is restricted to its integration into cellular environments. IEEE 802.21 provides the missing, technology-independent, abstraction layer able to offer a common interface to upper layers, thus hiding technology-specific primitives. This abstraction can be exploited by the IP stack (or any other upper layer) to better interact with the underlying technologies, ultimately leading to improved handover performance [2].

To achieve the aforementioned goals, 802.21 defines a media-independent entity that provides a generic interface between the different link layer technologies and the upper layers. To handle the particularities of each technology, 802.21 maps this generic interface to a set of media-dependent service access points (SAPs) whose aim is to collect information and control link behavior during handovers. In addition, a set of remote interfaces, terminal–network and network–network, are defined to convey the information stored at the operator's network to the appropriate locations (e.g. to assist the terminal in handover decisions). All of these aspects are covered by the 802.21 reference model and architecture, which are further analyzed in the following sections of this chapter. All the functionality of 802.21 is provided to the users by a set of services: Event, Command, and Information. These services are the core of the specification, and define the semantic model of the communication with the lower layers and the network. A detailed explanation of the services can be found in section 2.3.

The contribution of the 802.21 standard is centered on the following three main elements:
a) A framework that enables seamless handover between heterogeneous technologies. This framework is based on a protocol stack implemented in all the devices involved in the handover. The defined protocol stack aims to provide the necessary interactions among devices for optimizing handover decisions.
b) The definition of a new link layer SAP that offers a common interface for link layer functions and is independent of the technology specifics. For each of the technologies considered in 802.21, this SAP is mapped to the corresponding technology-specific primitives.
c) The definition of a set of handover enabling functions that provide the upper layers (e.g., mobility management protocols such as Mobile IP), with the required functionality to perform enhanced handovers. These functions trigger, via the 802.21 framework, that is the corresponding local or remote link layer primitives.

Although the main purpose of IEEE 802.21 is to enable handover between heterogeneous technologies, a set of secondary goals have also been defined. These secondary goals are [2]:
a) Service continuity: defined as the continuation of the service during and after the handover procedure. One of the main goals of 802.21 is to avoid the need to restart a session after a handover.
b) Handover-aware applications: The 802.21 framework provides applications with functions for participating in handover decisions. For instance, a voice application may decide to execute a handover during a silence period in order to minimize service disruption.
c) QoS-aware handovers: The 802.21 framework provides the necessary functions in order to make handover decisions based on QoS criteria. For instance, we may decide to hand over to a new network that guarantees the desired QoS.
d) Network discovery: This is an 802.21 feature that allows users to be provided with information on candidate neighbours for a handover.
e) Network selection assistance: Network selection is the process of making a handover decision based on several factors (e.g., QoS, throughput, policies, and billing). In line with the above, the 802.21 framework only provides the necessary functions to assist network selection, but does not make handover decisions, which are left to the higher layers.
f) Power management: For instance, power consumption can be minimized if the user is informed of network coverage maps, optimal link parameters, or sleep or idle modes.

2.2 GENERAL ARCHITECTURE

This standard supports different handover methods. Such methods are generally classified as “hard” or “soft”, depending on whether the handover procedure is “break-before-make” or “make-before-break” with respect to the data transport facilities that support the exchange of data packets between the MN and the network.

Handover decision making may involve cooperative use of both MN and network infrastructure.

2.2.1 MIHF Design Principles

This standard is based on the following general design principles:

a) MIHF is a logical entity that facilitates handover decision making. MIH Users make handover decisions based on inputs from the MIHF.

b) MIHF provides abstracted services to higher layers. The service primitives defined by this interface are based on the technology-specific protocol entities of the different access networks. The MIHF communicates with the lower layers of the mobility-management protocol stack through technology- specific interfaces.

c) Higher layer mobility management protocols specify handover signalling mechanisms for vertical handovers. Additionally, different access network technologies have defined handover signalling mechanisms to facilitate horizontal handover. The role of this standard is to serve as a handover facilitating service and to maximize the efficiency of such handovers by providing appropriate link layer intelligence and network information.

d) The standard provides support for remote events. Events are advisory in nature. The decision whether to cause a handover or not based on these events is outside the scope of this standard.

e) The standard supports transparent operation with legacy equipment. IEEE 802.21 standard compatible equipment should be able to co-exist with legacy equipment.

2.2.2 MIHF Services Overview

802.21 defines services that facilitate handovers between heterogeneous access links. These services are distinguished in the categories presented below:

1) A Media Independent Event Service (MIES) which provides event classification, event filtering and event reporting corresponding to dynamic changes in link characteristics, link status, and link quality.

2) A Media Independent Command Service (MICS) which enables MIH Users to manage and control link behavior relevant to handovers and mobility.

3) A Media Independent Information Service (MIIS) which provides details on the characteristics and services provided by the serving and neighbouring networks. The information enables effective system access and effective handover decisions.

The MIHF provides asynchronous and synchronous services through well-defined SAPs for link layers and MIH Users. In the case of a system with multiple network interfaces of arbitrary type, the MIH Users may use the Event Service, Command Service and Information Service provided by MIHF to manage, determine, and control the state of the underlying interfaces.

These services provided by MIHF help the MIH Users in maintaining service continuity, service adaptation to varying quality of service, battery life conservation, network discovery, and link selection. In a system containing heterogeneous network interfaces of IEEE 802 types and cellular (3GPP, 3GPP2) types, the MIHF may help the MIH Users to implement effective procedures to couple services across heterogeneous network interfaces. MIH Users may utilize services provided by the MIHF across different entities to query resources required for a handover operation between heterogeneous networks. MIH services in mobile nodes facilitate seamless handovers between heterogeneous networks. MIH Services are used by MIH Users such as a mobility management protocol (e.g., Mobile IP). Other mobility management protocols (in addition to Mobile IP) and even other MIH Users are not precluded from making use of MIH services.

2.2.3 MIHF Communication Model

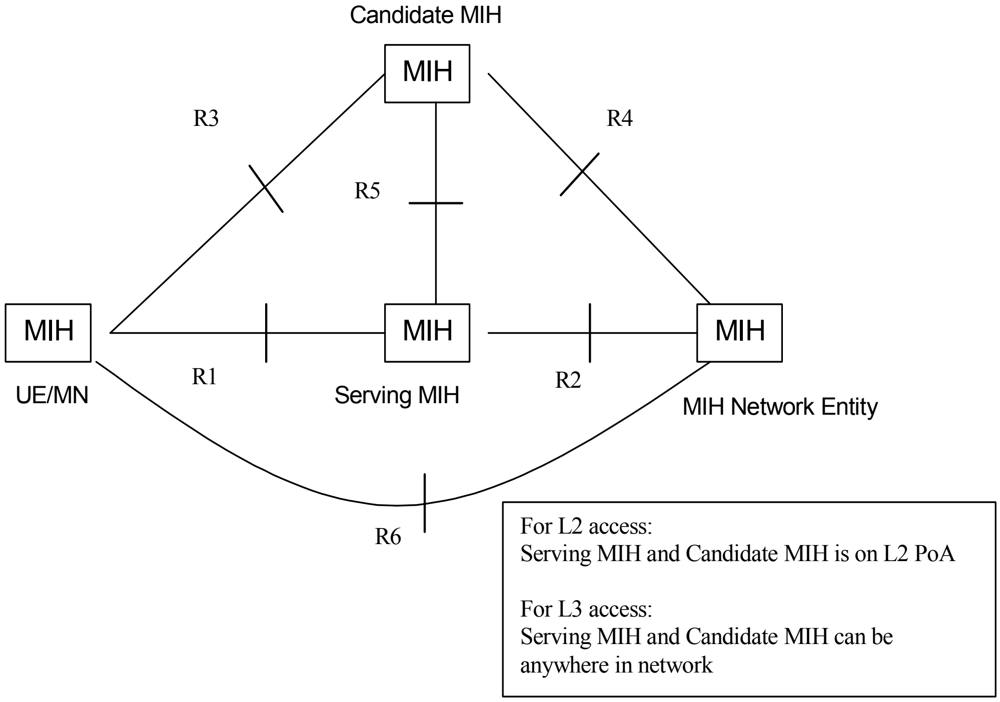


Figure 2.1: MIHF Communication Model [4]

MIHF functions communicate with each other for various purposes. The MN exchanges MIH information with its MIH Point of Service. The MIHF in any Network Entity becomes a MIH PoS when it communicates directly with an MN based MIHF. A MIH Network Entity may not have a direct connection to the MN and therefore does not constitute a MIH PoS for that particular MN. The same MIH Network Entity may still act as MIH PoS for a different MN. MIHF communication might not take place on all L2 interfaces of an MIH-capable MN. As an example, on a MIH-capable MN with three L2 interfaces namely IEEE 802.11, IEEE 802.16, and IEEE 802.3, the IEEE 802.3 interface might be used only for system administration and maintenance operations, while the IEEE 802.11 and IEEE 802.16 interfaces might engage in the provisioning of MIHF services. The MN may use L2 transport for exchanging MIH information with a MIH PoS that resides in the same Network Entity as its Network PoA. The MN may use L3 transport for exchanging MIH information with a MIH PoS that may not reside in the same Network Entity as its Network PoA.

The framework supports use of either L2 or L3 mechanisms for communication among MIH network entities.

Figure 2.1 shows the MIHF communication model. The model shows MIHFs in different distinctive roles and the communication relationships amongst them. The communication relationship shown in this figure applies only to MIHFs. It is important to note that each of the communication relationships in the communication model does not imply a particular transport mechanism. Rather, a communication relationship only intends to show that passing MIHF related information is possible between the two distinctive MIHFs. Moreover, each communication relationship shown in the diagram may encompass different types of interfaces, different transport mechanisms used (e.g., L2, L3), and different MIHF service related content being passed (e.g., MIIS, MICS, or MIES).

The communication model assigns different roles to the MIHF depending on its position in the system: a) MIHF on the MN, b) MIH PoS on the Network Entity that includes the serving PoA of the MN, c) MIH PoS on the Network Entity that includes a candidate PoA for the MN, d) MIH PoS on a Network Entity that does not include a PoA for the MN, e) MIH non-PoS on a Network Entity that does not include a PoA for the MN.

The communication model also identifies the following reference points (shown in figure 2.1) between different instances of MIHFs:
a) Reference point RP1: It refers to MIHF procedures between the MIHF on the MN and the MIH PoS on the Network Entity of its serving PoA. RP1 may encompass communication interfaces over both L2 and L3 and above. MIHF content passed over RP1 (as well as all the other RPs) may be related to MIIS, MIES, or MICS.
b) Reference point RP2: It refers to MIHF procedures between the MIHF on the MN and the MIH PoS on the Network Entity of a candidate PoA. RP2 may encompass communication interfaces over both L2 and L3 and above.
c) Reference point RP3: It refers to MIHF procedures between the MIHF on the MN and the MIH PoS on a non-PoA Network Entity. RP3 may encompass communication interfaces over L3 and above and possibly L2 transport protocols like Ethernet bridging, or MPLS.
d) Reference point RP4: It refers to MIHF procedures between an MIH PoS in a Network Entity and an MIH non-PoS instance in another Network Entity. RP4 may encompass communication interfaces over L3 and above.
e) Reference point RP5: It refers to MIHF procedures between two MIH PoS instances in distinct Network Entities. RP5 may encompass communication interfaces over L3 and above.

2.3 MIHF SERVICES

As we have already referred in the previous section, 802.21 defines three different types of communications with different associated semantics, the so-called MIH services: Event services, Command services and Information services. These services allow MIHF users to access handover-related information as well as deliver commands to the link layers or network. MIH services can be delivered asynchronously or synchronously. Events generated in link layers and transmitted to the MIHF or MIHF users are delivered asynchronously, while commands and information generated by a query/response mechanism are delivered synchronously [2].

2.3.1 Media Independent Event Service (MIES)

In general, handovers may be initiated either by the MN or by the network. Events that may be relevant to handover may originate from MAC, PHY or MIHF at the MN, at the network PoA, or at the PoS. Thus, the source of these events may be either local or remote. A transport protocol is needed for supporting remote events. Security is another important consideration in such transport protocols.

From the recipient's perspective the events are mostly advisory and not mandatory. In other words, the recipient is not obligated to act on the events, which are taking place in the lower layers. Layer 3 and above entities may also need to deal with reliability and robustness issues associated with these events. Higher layer protocols and other entities may prefer to take a more cautious approach when events originate remotely as opposed to when they originate locally. These events may also be used for horizontal handovers.

The Event Service is broadly divided into two categories, Link Events and MIH Events. In the following subsections, these two categories will be further analyzed. Both Link and MIH Events may traverse from a lower to a higher layer. Link Events are defined as events that originate from event source entities below the MIHF and may terminate at the MIHF. Entities generating Link Events include, but are not limited to, various IEEE 802-defined, 3GPP-defined, and 3GPP2-defined interfaces. Within the MIHF, Link Events may be further propagated, with or without additional processing, to MIH Users that have subscribed for the specific events. MIH events are defined as events created within the MIHF, or Link Events that are propagated by the MIHF to the MIH Users.

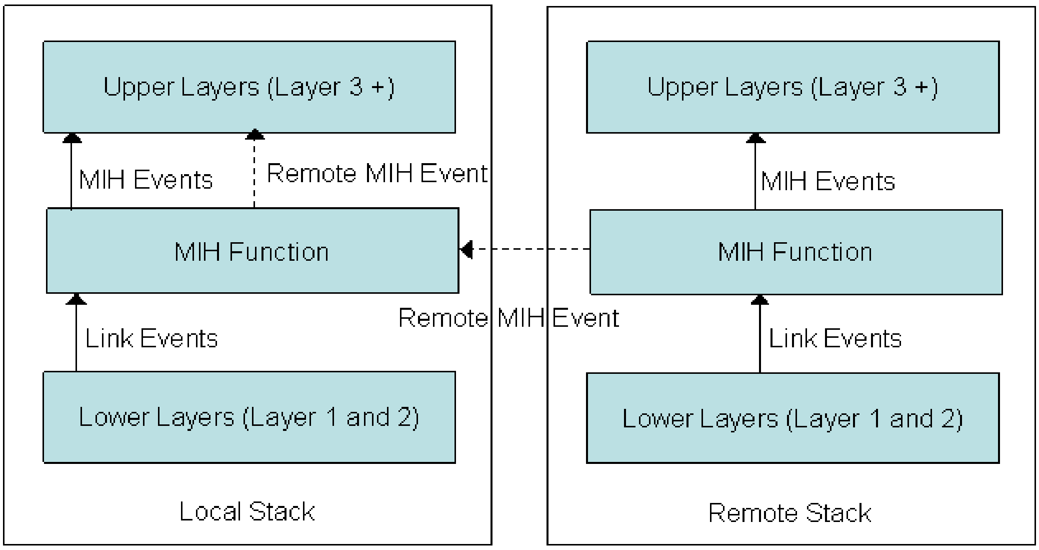


Figure 2.2: Remote MIH and Link events [5]

An event can be local or remote; a local event is one that propagates across different layers within the local protocol stack of a MIH entity, while a remote event is one that traverses across the network medium from one MIH entity to another MIH entity. All Link Events are local in nature and propagate from the local lower layer to the local MIHF. MIH Events may be local or remote. A remote MIH Event traverses the medium from a remote MIHF to the local MIHF and is then dispatched to local MIH Users that have subscribed to this remote event, as shown in Figure 2.2.

2.3.1.1 Link Events

The Media Independent Event Service supports several categories of link events:

a) MAC and PHY State Change events: These events correspond to changes in MAC and PHY state. For example Link_Up event is an example of a state change event.

b) Link Parameter events: These events are due to changes in link layer parameters. For example, the primitive Link_Parameters_Report is a Link Parameter event.

c) Predictive events: They convey the likelihood of a change in the link properties in the near future based on past and present conditions. For example, decay in signal strength of a WLAN network may indicate a loss of link connectivity in the near future. In case predictive events are incorrect they may be retracted.

d) Link Handover events: These events inform upper layers about the occurrence of L2 handovers/link switches if supported by the given media type.

e) Link Transmission events: These events indicate the link layer transmission status (e.g. success or failure) of upper layer PDUs. This information may be used by upper layers to improve buffer management for minimizing the upper layer data loss due to a handover. For example, the occurrence of a handover of a MN from one access network to another will result in the tear-down of the old link layer connection between the MN and the source access network and the establishment of a new link layer connection between the MN and the target access network. When this occurs, some upper layer PDUs may become “trapped” at the old link - including PDUs that had been queued at the old link but never been transmitted before the link was torn-down (i.e., unsent PDUs), and PDUs that have been transmitted over the old link but never been fully acknowledged by the upper layer receiver before the link was torn down (i.e., unacked PDUs). These “trapped” PDUs will be discarded when the old link is torn down. As a result, unless the upper layer sender attempts to retransmit them over the new link connection, these upper layer PDUs will never reach the receiver. However, before the retransmission can happen, the upper layer sender first needs to be notified about which upper layer PDUs are “trapped” at the old link. Link Transmission events facilitate this process by providing the upper layer a local indication on whether a particular upper layer PDU has been successfully transmitted or not.

In general when a link event occurs due to a change in link condition, it is not known at that instant if this would lead to intra-technology handover or inter-technology handover. That determination is done higher up in the protocol stack by the network selection entity based on variety of other factors. As such certain link layer events such as Link_Going_Down may lead to either intra-technology or inter technology handovers. The network selection entity may try to maintain the current connection, by first trying intra-technology handovers and only later on resort to inter-technology handovers.

2.3.1.2 MIH Events

In this category we have a list of events that will be further analyzed in section 2.4. In the current subsection, we simply refer these events in a sententious manner. The MIH_Link_Detected event is typically generated on the MN when the first PoA of an access network is detected. This event is not generated when subsequent PoAs of the same access network are discovered. With MIH_Link_Up event, L2 connection is established and link is available for use. The opposite event is MIH_Link_Down, where L2 connection is broken and link is not available for use. MIH_Link_Event_Rollback is triggered when previous link event needs to be rolled back. When MIH_Link_Parameters_Report is triggered, link parameters have crossed a specified threshold and need to be reported. If link conditions are degrading and connection loss is imminent, the MIH_Link_Going_Down event takes place. The transmission of MIH_Link_Handover_Imminent takes place, whenever a L2 handover is imminent based on either the changes in the link conditions or additional information available in the network. For example, the network may decide that an application requires a specific QoS that can be best provided by a certain access technology. The use of MIH_Link_Handover_Complete and MIH_Link_ PDU_Transmit_Status events is obvious.

2.3.2 Media Independent Command Service (MICS)

The media-independent command service (MICS) refers to the commands sent from the higher layers to the lower layers in order to determine the status of links or control and configure the terminal to gain optimal performance or facilitate optimal handover policies. The mobility management protocols should combine dynamic information regarding link status and parameters, provided by the MICS with static information regarding network status, network operators, or higher-layer service information provided by the media-independent information service, to help in the decision making. The receipt of a certain command request may cause event generation, and in this way the consequences of a command could be followed by the network and related entities. Commands can be delivered locally or remotely. Through remote commands the network may force a terminal to hand over, allowing the use of network initiated handovers and network assisted handovers [2]. A set of commands are defined in the specification to allow the user to control lower layers configuration and behavior, and to this end some PHY layer commands have being specified too. The Command Service is broadly divided into two categories (as Event Service), Link Commands and MIH Commands. In the following subsections, these two categories will be further analyzed.

Remote MIH commands are sent by MIH Users to the MIHF in a peer protocol stack. A remote MIH command delivered to a peer MIHF is executed by the lower-layers under the peer MIHF as a link command; or is executed by the peer MIHF itself as a MIH command (as if the MIH command came from a MIH User of the peer MIHF); or is executed by a MIH User of the peer MIHF in response to the corresponding indication. Often, a MIH indication to a remote MIH User may result from the execution of the MIH command by the peer MIHF. Figure 2.3 shows remote MIH commands.

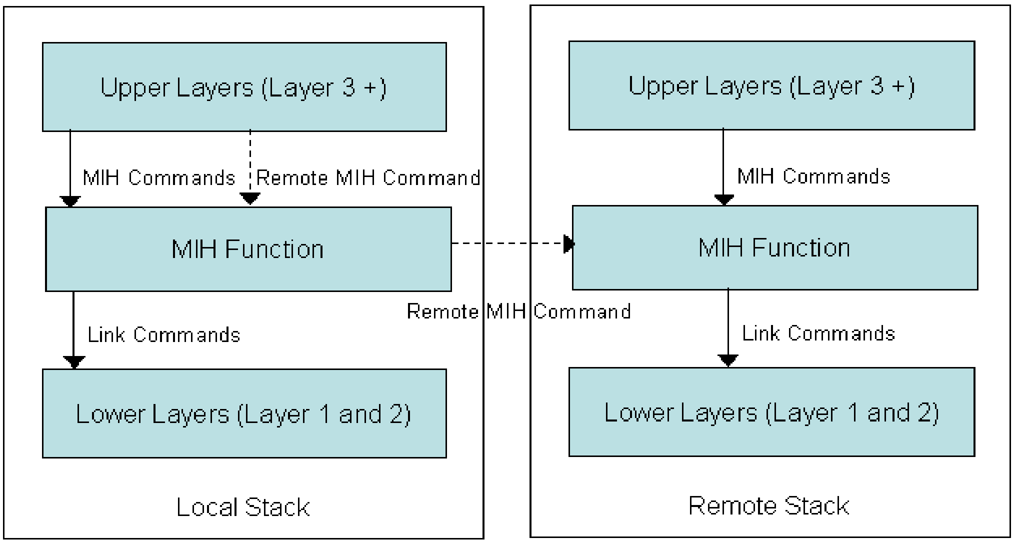


Figure 2.3: Remote MIH and Link Commands [5]

2.3.2.1 Link Commands

There is a list of six main link commands. Link_Capability_Discover command queries and discovers the list of supported link layer events and link layer commands. Triggering Link_Event_Subscribe command, a subscription to one or more events from a link takes place. Link_Event_Unsubscribe is the exact opposite command. With Link_Get_Parameters command, we get parameters measured by the active link, such as signal-to-noise ratio, BER, RSSI etc. A configuration of thresholds for Link Parameters Report event is taking place with the assistance of Link_Configure_Thresholds command. Eventually, with Link_Action command we request an action on a link layer connection. Details about all these commands will be demonstrated in the next section.

2.3.2.2 MIH Commands

Generally, there are three types of MIH handover command primitives based on the functionality specified for the following scenarios: i) MN to Network (MIH_MN_HO_ prefix), ii) Network to MN (MIH_Net_HO_ prefix), and iii) Network to Network (MIH_N2N_HO_ prefix). This classification helps to ensure the specification of right protocol functionality and the relevant parameters for specific use as determined by the origination and the destination points.

MIH_Link_Get_Parameters command gets the status of a link. MIH_Link_Configure_Thresholds configures link parameter thresholds. MIH_Link_Actions controls the behaviour of a set of links. In MIH_Net_HO_Candidate_Query, the network may initiate handover and send a list of suggested networks and associated PoAs. MIH_MN_HO_Candidate_Query command is used by MN to query and obtain handover related information about possible candidate networks. MIH_N2N_HO_Query_Resources is sent by the serving MIHF entity to the target MIHF entity to allow for resource query. In case of MIH_Net_HO_Commit, the network commits to do the handover and sends the choice of selected network and associated PoA. MIH_N2N_HO_Commit command is used by a serving network to inform a target network that a MN is about to move toward that network, initiate context transfer (if applicable) and perform handover preparation. Eventually, MIH_MN_HO_Complete and MIH_N2N_HO_Complete commands provide a notification indicating the status of the handover completion.

2.3.3 Media Independent Information Service (MIIS)

Media Independent Information Service (MIIS) provides a framework by which a MIHF, residing in the MN or in the network, may discover and obtain network information within a geographical area to facilitate network selection and handovers. The objective is to acquire a global view of all the heterogeneous networks relevant to the MN in the area to facilitate seamless roaming across these networks. Media Independent Information Service includes support for various Information Elements (IEs). Information Elements provide information that is essential for a network selector to make intelligent handover decisions. Depending on the type of mobility, support for different types of information elements may be necessary for performing handovers. MIIS provides the capability for obtaining information about lower layers such as neighbor maps and other link layer parameters, as well as information about available higher layer services such as internet connectivity.

MIIS provides a generic mechanism to allow a service provider and a mobile user to exchange information on different handover candidate access networks. The handover candidate information may include different access technologies such as IEEE 802 networks, 3GPP and 3GPP2 networks. The MIIS also allows this collective information to be accessed from any single network. For example, by using an IEEE 802.11 access network it may be possible to get information not only about all other IEEE 802 based networks in a particular region but also about 3GPP and 3GPP2 networks. Similarly using a 3GPP2 interface, it may be possible to get access to information about all IEEE 802 and 3GPP networks in a given region. This capability allows the MN to use its currently active access network and inquire about other available access networks in a geographical region. Thus, a MN is freed from the burden of powering up each of its individual radios and establishing network connectivity for the purpose of retrieving heterogeneous network information. MIIS enables this functionality across all available access networks by providing a uniform way to retrieve heterogeneous network information in any geographical area.

The main goal behind the Information Service is to allow MN and network entities to discover information that may influence the selection of appropriate networks during handovers. This information is intended to be primarily used by a policy engine entity that may make effective handover decisions based on this information. This Information Service provides mostly static information, although network configuration changes are also accounted for. Other dynamic information about different access networks, such as current available resource levels, state parameters, and dynamic statistics should be obtained directly from the respective access networks. Some of the key motivations behind the Information Service are as follows:

a) Provide information about the availability of access networks in a geographical area.
b) Provide static link layer information parameters that could help the mobile nodes in selecting the appropriate access network.
c) Provide information about capabilities of different PoAs in neighbour reports to aid in configuring the radios optimally (to the extent possible) for connecting to available/selected access networks. However, for the most part, dynamic link layer parameters have to be obtained or selected based on direct interaction with the access networks as the Information Service may not be able to help much in this regard.
d) Provide an indication of higher layer services supported by different access networks and core networks that may aid in making handover decisions. Such information may not be available (or could not be made available) directly from the MAC sublayer or PHY of specific access networks, but could be provided as part of the Information Service.

2.3.3.1 Information Elements (IEs)

The information elements (IE) provided by the MIIS can be divided in the following groups [2]:
a) General information: These IEs give a general overview about the networks covering a specific area such as network type, operator identifier, or service provider identifier (IE_NETWORK_TYPE, IE_OPERATOR_ID, IE_COUNTRY_CODE, IE_SERVICE_PROVIDER_ID).
b) Access network-specific information: These IEs provide specific information for each technology and operator. The information is related to security characteristics, QoS information, revisions of the current technology standard in use, cost, roaming partners etc (IE_COST, IE_ROAMING_PARTNERS, IE_NETWORK_ID, IE_NETWORK_QOS, IE_NETWORK_DATA_RATE, IE_NET_CAPABILITIES, IE_NET_SUPPORTED_LCP etc).
c) Point of attachment specific information: These IEs provide information for each PoA (for each technology and operator). The information comprises aspects like MAC address of the PoA, geographical location, data rate, channel range, and so on (IE_POA_MAC_ADDR, IE_POA_LOCATION, IE_POA_CHANNEL_RANGE, IE_POA_SYSTEM_INFO).
d) PoA specific higher-layer service: The information provided is related to the available services on this PoA and network. The information provided may be the number of subnets this PoS supports, the IP configuration methods available, or even a list of all supported services of the PoA (IE_POA_SUBNET_INFO, IE_POA_IP_ADDR).
e) Other information can be added, such as vendor- specific information or services.

Concluding this subsection, it is important to note that the mobile node should be able to discover whether the network supports IEEE 802.21 by use of a discovery mechanism or information obtained by MIIS through another interface. It is also important that the mobile terminal is able to obtain MIIS information even before the authentication in the PoA is performed in order to be able to check the security protocols, support of QoS, or other parameters before performing a handover. The communication between the different entities of the IEEE 802.21 network in order to gather information related to the MIIS may be performed through all the communication reference points defined in subsection 2.2.3.

2.4 MIH SERVICE ACCESS POINTS (MIH SAPs)

The MIH Function uses the following SAPs for interfacing with other entities:
a) MIH_LINK_SAP: Abstract media dependent interface of MIHF with the lower layers of the media-specific protocol stacks. For more details see subsection 2.4.1.
b) MIH_SAP: This SAP defines the media independent interface between the MIHF and MIH Users. For more details see subsection 2.4.2.
c) MIH_NET_SAP: Abstract media dependent interface of MIHF which provides transport services over the data plane on the local node, supporting the exchange of MIH information and messages with the remote MIHF. For more details see subsection 2.4.3.

2.4.1 MIH_LINK_SAP primitives

There are plenty of primitives defined in 802.21 standard belonging to this category. We will particularly focus on the most important ones, in order to understand the main concept of the information exchanged between the mobile terminals and the network entities. Thus, the reader will be able to be familiar with the use cases demonstrated in the next chapter. For each primitive, we first analyze its function, we refer its trigger for generation and finally we investigate its effect on receipt.

2.4.1.1 Link_Detected.indication

Link_Detected primitive indicates the presence of a new PoA. This may imply that the MN is in the coverage area and may listen to a beacon frame, or that the MN may have received a response to a probe. Link_Detected does not guarantee that the MN would be able to establish connectivity with the detected link, but just that the MN may attempt to gain connectivity. MIH Users and the MIHF itself may evaluate additional properties of the link before attempting to establish a L2 connection with the link. Moreover, Link_Detected is not generated when additional PoAs of the same link are discovered. In case of 802.11, this is generated by MSCGF.

This type of message contains parameters such as link and access network identifier, signal strength and data rate. It is generated on the MN when the first PoA of an access network is detected. This event is not generated when subsequent PoAs of the same access network are discovered during the active connection on that link. The MIHF receives this event from the link layer. The MIHF then passes this notification to the MIH User(s) which have subscribed for this notification. The MIH User(s), including the MIHF itself, may discover additional properties of the link before selecting it for establishing connectivity.

2.4.1.2 Link_Parameters_Report.indication

Link_Parameters_Report indicates changes in link parameters that have crossed specified threshold levels. Link_Parameters_Report may also be generated at specified intervals for various parameters. The parameters of this primitive are just the link identifier and a list of link parameters report. For each specified parameter, this notification is generated either at a predefined regular interval determined by a user configurable timer or when it crosses a configured threshold. The MIHF receives this event from the link layer. The MIHF then passes this notification to the MIH User(s) which have subscribed for this notification. The MIH User(s) may take different actions on this notification. If parameters related to link quality cross a certain threshold then that link may need to be evaluated for handing over current connections. The MIHF may collectively evaluate different parameters and give appropriate indications to higher layers regarding suitability of different links.

2.4.1.3 Link_PDU_Transmit_Status.indication

Link_PDU_Transmit_Status indicates the transmission status of a higher layer PDU by the link layer. A success status indicates that the higher layer PDU has been successfully delivered from the link layer in the local node to the link layer in the peer node. A higher layer intermediate buffer management entity could use this indication to flush the delivered PDU from its buffer. A failure status indicates that the higher layer PDU identified in the indication was not delivered successfully from the link layer in the local node to the link layer in the peer node. During a handover, if such a failure indication is received from the link connection with the source network, the higher layer intermediate buffer management entity could attempt to retransmit the failed PDU once a connection to the target network is established. A Packet Identifier is expected to be passed alongside when each higher layer PDU is sent from the higher layer to the link for transmission. The Packet Identifier is defined in this standard as a container structure whose syntax and semantics will be decided by the upper layer (i.e., the MIH User which subscribes to this event). The MIHF and link layer just pass and return the Packet Identifier and do not need to understand its syntax and semantics. To avoid receiving excessive amount of link PDU transmission status indications, an MIH User, for example, may choose to subscribe to this event only after it receives a Link_Handover_Imminent.indication or when it is about to invoke an MIH_Link_Actions.request to perform a handover, and to unsubscribe from the event once it receives indication that the handover is completed.

2.4.1.4 Link_Get_Parameters.request/confirm

This primitive is used by the MIHF to obtain the current value of a set of link parameters of a specific link. The parameters of this message are LinkParameters-Request, LinkStatesRequest and LinkDescriptorsRequest. The recipient link responds with Link_Get_Parameters.confirm primitive. This primitive provides current value of the requested link parameters and has the same parameters with the corresponding request message. The recipient may pass the link parameter values received to the MIH Users. However, if Status does not indicate "Success", the recipient performs appropriate error handling.

2.4.1.5 Link_Action.request/confirm

The Link_Action.request primitive is used by the MIHF to request an action on a link layer connection to enable optimal handling of link layer resources for the purpose of handovers. The link layer connection can be ordered, e.g. to shut down, to remain active, to perform a scan, or to come up active and remain in stand-by mode. The command execution delay time can also be specified for cases where the link layer technology under consideration supports the action. The MIHF generates this primitive upon request from the MIH User to perform an action on a pre-defined link layer connection. Upon receipt of this primitive, the link layer technology supporting the current link layer connections performs the action specified by the Link Action parameter in accordance with the procedures specified by the relevant standards organization and at the time specified by the execution delay parameter. After these procedures, the link layer technology generates the Link_Action.confirm primitive in order to communicate the result of the action executed on the link layer connection. Upon receipt of this primitive, the MIHF determines the relevant MIH command that needs to be used to provide an indication or confirmation to the MIH User of the actions performed on the current link layer connection.

2.4.2 MIH_SAP primitives

Following the same way in distributing the most important primitives of this category we have selected the following eight ones:

2.4.2.1 MIH_Link_Detected.indication

The MIH_Link_Detected.indication is sent to local MIHF users to notify them of a local event (i.e., Link_Detected.indication described in 2.4.1.1), or is the result of the receipt of an MIH_Link_Detected indication message to indicate to the remote MIHF users, who have subscribed to this remote event, that a remote Link_Detected event occurred. The parameters of this message are just the Source Identifier and the Link Detected Info List. The effect on receipt of this message is MIH user dependant.

2.4.2.2 MIH_Link_Parameters_Report.indication

MIH_Link_Parameters_Report indication is sent by the local MIHF to a local MIH User to report the status of a set of parameters of a local or remote link. This notification is generated by the local MIHF either: a) at a predefined regular interval determined by a user configurable timer, b) when a specified parameter of a currently active local interface crosses a configured threshold. In such a case, the local MIHF most likely will first receive a Link_Parameters_Report.indication fromthe local link layer, c) when a MIH_Link_Parameters_Report indication message is received from a remote MIHF. Upper layer entities may take different actions upon receipt of this indication.

2.4.2.3 MIH_Link_Get_Parameters.request/confirm

In general, a MIH_Link_Get_Parameters command is issued by upper layer entities to discover and monitor the status of the currently connected and potentially available links. The destination of a MIH_Link_Get_Parameters command may be local or remote. For example, a MIH_Link_Get_Parameters request issued by a local upper layer may help the policy function that resides out of the MIH to make optimal handover decisions for different applications when multiple links are available in a MN. However, a remotely initiated MIH_Link_Get_Parameters request from the network side may enable the network to collect the status information on multiple links in a MN through the currently connected link.

More specifically, the MIH_Link_Get_Parameters.request primitive is invoked by a MIH User when it wants to request the status information of a set of local or remote links. If the destination of the request is the local MIHF itself, the local MIHF gets the requested information on the status of the specified local links and responds with a MIH_Link_Get_Parameters.confirm. If the destination of the request is a remote MIHF, the local MIHF generates and sends a MIH_Link_Get_Parameters request message to the remote MIHF. The confirmation message returns the results of a MIH_Link_Get_Parameters request to the requesting MIH User. Upon receipt of the link status information, the MIH User makes appropriate decisions and takes suitable actions. However, if status does not indicate “Success”, the recipient performs appropriate error handling.

2.4.2.4 MIH_Link_Actions.request/confirm

The request message is invoked by a MIH User when it attempts to control the behavior of a set of local or remote lower layer links. If the destination of the request is the local MIHF itself, the local MIHF issues Link_Action request(s) to the specified lower layer link(s). If the destination of the request is a remote MIHF, the local MIHF generates and sends a MIH_Link_Actions request message to the remote MIHF. Upon the receipt of the message, the remote MIHF then issues a Link_Action request(s) to the specified lower layer link(s). The confirmation message returns the result of a MIH_Link_Actions.request to the requesting MIH User. The parameters of this message are, correspondingly to the request message, the Source Identifier and the Link Actions Result List. Upon receipt of the result, the MIH User makes appropriate evaluations and takes any suitable actions. However, if Status does not indicate “Success”, the recipient performs appropriate error handling.

2.4.2.5 MIH_Net_HO_Candidate_Query / MIH_MN_HO_Candidate_Query

For network initiated handovers, the network controller provides a list of candidate network choices to the MN (via MIH_Net_HO_Candidate_Query request message). The MN indicates resources required on each of these candidate networks in the MIH_Net_HO_Candidate_Query response message. The network controller may then query each of the candidate networks for available resources (using MIH_N2N_HO_Query_Resources primitive). Once the target network has been selected, the network controller may send a MIH_Net_HO_Commit message. An example of this operation is illustrated in the next chapter (section 3.1). Conclusively, there are three types of messages supporting these messages; that is indication, response and confirmation messages.

The MIH_MN_HO_Candidate_Query primitive is used by MIH Users on a MN to inform MIHF to query candidates for possible handover initiation. The request may include queries on QoS resources and/or whether IP address configuration method of the ongoing data sessions can be supported in the candidate network. The main parameters of this message are Destination Identifier, Source Link Identifier, Candidate Link List and Query Resource List. This primitive also includes the current IP configuration server address (e.g., DHCP server, FA IP address, AR IP address) when the current IP configuration method is included. It is generated by a MIH User in the MN that may want to query other candidate networks for a possible handover. MN uses the Query Resource List parameter to notify the serving PoS of the minimal resource requirement at the candidate networks in order for the handover to be successful. A MIH User on MN may generate this primitive when it wants to query IP address related information from the candidate networks before handover. Upon receipt of this primitive, the local MIHF generates and sends a MIH_MN_HO_Candidate_Query request message to the remote MIHF identified by the Destination Identifier.

2.4.2.6 MIH_N2N_HO_Query_Resources

The MIH_N2N_HO_Query_Resources.request primitive is used by a MIHF on the serving network to communicate with its peer MIHF on the candidate network. This is used to query the available link resource and IP address related information of the candidate network. It is generated after receiving the MIH_MN_HO_Candidate_Query request message from the MIHF on the MN in the case of mobile-initiated handover (see subsection 3.1.1). In the case of network-initiated handover (see subsection 3.1.2), this primitive is generated after receiving the MIH_Net_HO_Candidate_Query response message from the MN. Upon receipt of this primitive MIHF sends a request message to the destination.

With MIH_N2N_HO_Query_Resources.indication primitive, the MIHF on the candidate network indicates that a MIH_N2N_HO_Query_Resources.request message is received from a remote MIHF on the serving network so that the upper layer entity can identify the link resource usage and provide IP address related information for the impending handover. It is generated by MIHF when the MIHF on the candidate network receives MIH_N2N_HO_Query_Resources request message from a peer MIHF on the serving network. The MIH User on the candidate network identifies the link resource usage for the impending handover. It also replies with MIH_N2N_HO_Query_Resources.response primitive.

MIH_N2N_HO_Query_Resources.response primitive is used by a MIHF on the candidate network to communicate with its peer MIHF on the serving network which sent out a MIH_N2N_HO_Query_Resources request. This is used to notify the MIHF on the serving network of the link resource status of the candidate network. It is also used to provide IP address related information of the candidate networks. The MIHF on the candidate network invokes this primitive in response to a

MIH_N2N_HO_Query_Resources.request message from a peer MIHF entity on the serving network. Upon receipt of this primitive MIHF sends a response message to the destination.

Eventually, MIH_N2N_HO_Query_Resources.confirm primitive is generated by the MIHF on the serving network to respond with the result of any resource preparation for the impending handover and to notify the link resource status of the candidate network. It also carries IP address related information on the candidate networks to MIH Users on the serving network.

2.4.2.7 MIH_Net_HO_Commit / MIH_N2N_HO_Commit

The MIH_Net_HO_Commit.request primitive is used by a MIH User on the network to communicate with the remote MIH User on the MN. The primitive is used to request the peer MIH User the commitment to perform a network-controlled or network-assisted link handover based on selected choices for candidate networks and PoA. Upon receipt of this primitive, MIHF sends a request message to the destination. MIH_Net_HO_Commit.indication primitive is generated by a MIHF to indicate that a MIH_Net_HO_Commit.request message has been received from a peer MIHF. The MIH User receiving this primitive replies with a MIH_Net_HO_Commit.response primitive. Only the applicable actions in the Link Actions List are executed. The non-applicable link actions indicate failed actions when preparing the response. MIH_Net_HO_Commit.response is used by a MIHF to communicate with a peer MIHF from which a MIH_Net_HO_Commit.request message is received. The primitive is used to communicate the response of a handover commit request. Upon receipt, the old Serving PoS is informed about the status of the previously issued command request. Since MIH_Net_HO_Commit contains actions to effect the handover, and the response has the status of those actions, the link to the old PoS may not be accessible (e.g. break before make) to receive the response before L3 connectivity has been established on the new link and only if the MN knows the old PoS L3 address thus, the old PoS may not receive this response. MIH_Net_HO_Commit.confirm is generated by the MIHF to confirm that a MIH_Net_HO_Commit.response message is received from a peer MIHF. Upon receipt, the old serving PoS is informed about the status of the previously issued command request.

MIH_N2N_HO_Commit.request primitive is used by a MIH User on the serving network to inform a selected target network that a MN is about to move to the target network. Upon receipt of this primitive, the local MIHF generates and sends a MIH_N2N_HO_Commit.request message to the remote MIHF on the selected target network identified by the Destination Identifier. MIH_N2N_HO_Commit.indication primitive is generated by a MIHF to indicate that a MIH_N2N_HO_Commit.request message has been received from a peer MIHF on serving network. Upon receipt of this primitive, MIH User generates a MIH_N2N_HO_Commit.response primitive. The latter is generated by a MIHF User in response to a received MIH_N2N_HO_Commit.indication primitive. Upon receipt, the MIHF generates and sends a MIH_N2N_HO_Commit.response message to the peer MIHF on the serving network which sent a MIH_N2N_HO_Commit.request message. MIH_N2N_HO_

Commit.confirm primitive is generated by the MIHF to confirm that a corresponding response message is received from a peer MIHF on the selected target network. Upon receipt, the serving network is informed about the status of the previously issued command request so that it can react accordingly. For instance, the serving network may determine that the handover procedure is acknowledged by the target network and it can notify the MN to perform handover. However, if Status does not indicate "Success", the recipient ignores any other returned values and, instead, performs appropriate error handling.

2.4.2.8 MIH_Get_Information

MIH_Get_Information.request primitive is used by a MIH User to request information from a MIH information server. The information query may be related to a specific interface, attributes to the network interface, as well as the entire network capability. The service primitive has the flexibility to query either a specific data within a network interface or extended schema of a given network. It is assumed that the available information could be broadcast in access technology specific manner such as in 802.11 and 802.16. The order of the queries in each of Info Query Binary Data List, Info Query RDF Data List and Info Query RDF Schema List parameters identifies the priority of the query. The first query has the highest priority to be processed by MIIS. The recipient will forward the query in a MIH message to the designated MIIS server. The MIIS server tries to interpret the query request and retrieve the specified information.

MIH_Get_Information.indication primitive is generated by the MIHF to indicate that a MIH_Get_Information Request message is received from a peer MIHF. The recipient tries to interpret the query request and retrieve the specified information. Once the information is retrieved, the recipient replies with the MIH_Get_Information .response primitive. The latter is generated by a MIH User in response to a received MIH_Get_Information.indication message. When the size of the Info Response parameters exceeds the maximum size specified in the Max Response Size parameter from the corresponding indication message, one or more of the lower order list elements in Info Response parameters may be omitted. The recipient will return a MIH_Get_Information.response message to the designated MIIS client. MIH_Get_Information.confirm is generated by the MIHF on receiving a MIH_Get_Information.response message from a peer MIHF. The MIH User that requested the information tries to interpret the Info Response parameters and takes suitable action.

2.4.3 MIH_NET_SAP primitives

The primitives associated with data transfers are as follows: a)MIH_TP_Data.Request, b) MIH_TP_Data.indication and c) MIH_TP_Data.confirm. The MIHF uses the first primitive to request that a MIH PDU be transported. The transport service provider uses the second primitive to indicate the arrival of a MIH PDU. Finally, the third primitive is used to acknowledge the successful transfer of the MIH PDU.

2.4.3.1 MIH_TP_Data

MIH_TP_Data.request primitive is used to request that a MIH PDU be transported to a remote MIHF. The receipt of this message causes the selected transport service provider to attempt to transport the MIH PDU. The parameters of this type of message are Transport Type, Source and Destination Address, Reliable Delivery Flag and MIH Protocol PDU. MIH_TP_Data.indication is used by the transport service provider to indicate that a MIH PDU has been received from a remote MIHF. The receipt of this message causes the MIHF to receive the MIH PDU that was transported. MIH_TP_Data.confirm is passed from the transport service provider to the MIHF to confirm that a request to transfer a MIH PDU succeeded. Upon receipt of this primitive, the receiving MIHF may stop its retransmission timer for the corresponding request. When the MIHF does not receive this primitive for a predefined time after transmitting a MIH_TP_Data.request with Reliable Delivery Flag set to TRUE, the MIHF may try to retransmit the MIH_TP_Data.request.

2.5. MIH PROTOCOL

The MIHF entities in MN and network entities communicate with each other using the MIH protocol messages specified in this section. The MIH protocol defines message formats for exchanging these messages between peer MIH Function entities. These messages are based on the primitives which are part of the MIH Services and have already been extensively analyzed in the previous section.

MIH protocol messages require reliability to ensure the receipt of data by the intended destination. Reliability may be provisioned with the optional acknowledgement service as part of the MIH protocol. The acknowledgement service is particularly useful when the underlying transport used for remote communication does not provide reliable services. When the MIH transport is reliable, the acknowledgement service is optional. When seeking acknowledgement service, the source MIH node shall start a retransmission timer after sending a MIH protocol message with the ACK-Req bit set and shall save a copy of the MIH protocol message while the timer is active. If the ACK message is not received before the expiration of the timer, the source MIH node may immediately retransmit the saved message with the same Message-ID and with the same Transaction-ID. If the source MIH node receives the acknowledgement before the expiration of the timer on the first or any subsequent retransmitted attempt, then the source MIH node has ensured the receipt of the MIH packet and therefore, may reset the timer and release the saved copy of the MIH protocol message. During retransmission, if the source MIH node receives the acknowledgement for any of the previous transmission attempts then the source MIH node may determine successful delivery of the message and may not have to wait for any further acknowledgements for the current message. All the above play a significant role in next generation wireless systems. Our ulterior goal is to define metrics for delay or jitter during real time delay sensitive services of the future such as streaming video, live video calls, VoIP calls etc.

The MIH protocol does not provide direct support for congestion control. Therefore, it is recommended to run the MIH protocol over congestion aware transport layers. In order to help prevent congestion, flow control mechanisms are implemented at the MIHF. A single rate limiter applies to all traffic (for all interfaces and message types). It applies to retransmissions, as well as new messages, although an implementation may choose to prioritize one over the other. When the rate limiter is in effect, MIH messages are queued until transmission is re-enabled, or an error condition may be indicated back to local upper layer applications. When a MIHF suffers from overload, it may drop requests from MIH requestors, and may differentiate between requestors to implement selective dropping. For example, messages could be dropped from a particular requestor if that requestor could be established as the origin of a denial of service attack. A MIHF may not drop a message if it was delivered reliably by the transport (L2 or L3). Any reliable delivery function may be able to indicate a flow control back to the requestor, and a MIHF may invoke flow control towards a specific requestor when overloaded with reliably delivered messages. MIH protocol messages are delivered via media dependent transport. The MIH Function determines the transport reliability provided to the MIH user. However, the MIH user may provide its preference to the MIH Function. How the preference is communicated to the MIHF depends on the implementation and we don’t really care about it. If the MIHF uses unreliable transport and a reliable message delivery is required, the MIH protocol ACK operation shall be applied for transmitting the message. If the MIHF uses reliable transport, then the MIH protocol ACK operation should not be required for transmitting the message.

2.6 REFERENCES

[1] George Lampropoulos, Apostolis Salkintzis and Nikos Passas, “**Media-Independent Handover for Seamless Service Provision in Heterogeneous Networks**”, IEEE Communications Magazine, January 2008.

[2] Antonio de la Oliva et al, “**An Overview of IEEE 802.21: Media-Independent Handover Services**”, IEEE Communications Magazine, August 2008.

[3] IEEE P802.21/D9.0, “**Draft Standard for Local and Metropolitan Area Networks: Media Independent Handover Services**”, February 2008.

[4] http://www.ieee802.org/21/doctree/2005_Meeting_Docs/2005-09_meeting_docs/21-05-0365-02-0000-communication_model_text_(draft).doc

[5] http://www.rajjain.com/cse574-06/ftp/handover/index.html

CHAPTER 3
USE CASES FOR HANDOVER PROCEDURES BASED ON 802.21

3.1 GENERAL HANDOVER PROCEDURE

Let us assume that there is a MN which has multiple interfaces in the 802.21 framework. When the channel quality of the MN is about to become poor in the current network, MIES delivers a triggered event to indicate the status of the physical and link layers. The events include Link Parameter Change, Link Down, Link Going Down, etc. If a communication protocol receives the Link Going Down event from the link layer through MIES, then it can gather the information of neighbour networks through MIIS to find the most proper network to handover. Next, the upper layer commands to conduct a handover to the selected network through MICS. Through this generally described process, the MN executes a handover based on the 802.21 framework [1].

In this section, we introduce the terminal-initiated handover and network-initiated handover which have been defined in the 802.21 standard. The main difference between the terminal-initiated handover and the network-initiated handover depends on which entity mainly controls handovers. The MN itself decides the execution of a handover in the former case, while the serving network of the MN decides the execution of handover in the latter case. All the messages that are referred in the next two subsections should be familiar to the reader. In a possible opposite case, the reader can find details in sections 2.3 and 2.4 of the present work.

3.1.1. Mobile-initiated Handover procedure

The Mobile-initiated Handover procedure operates as follows [2, 6]:
1) Mobile Node is connected to the serving network via Current PoS and it has access to MIH Information Server (MIIS).
2) Mobile Node queries information about neighbouring networks by sending the MIH_Get_Information.request to Information Server. Information Server responds with MIH_Get_Information.response. This information query may be attempted as soon as Mobile Node is first attached to the network.
3) Mobile Node triggers a mobile-initiated handover by sending MIH_MN_HO_Candidate_Query.request to Serving PoS. This request contains the information of potential candidate networks.
4) Serving PoS queries the availability of resources at the candidate networks by sending MIH_N2N_HO_Query_Resources.request to one or multiple candidate PoSs.
5) Candidate PoSs respond with MIH_N2N_HO_Query_Resources.response and Serving PoS notifies the Mobile Node of the resulting resource availability at the candidate networks through MIH_MN_HO_Candidate_Query.response.
6) Mobile Node decides the target of the handover and commits a link switch to the target network interface.

7) Serving PoS may send the MIH_N2N_HO_Commit.request to target PoS to request resource preparation at the target network. Target PoS responds the result of the resource preparation by MIH_N2N_HO_Commit.response.
8) The new layer 2 connection is established and certain mobility management protocol procedures are carried out between Mobile Node and target network.
9) Mobile Node may send MIH_MN_HO_Complete.request to Target PoS. Target PoS sends MIH_N2N_HO_Complete.request to previous Serving PoS to release resources which were allocated to Mobile Node. After identifying that resources are successfully released, target PoS may send MIH_MN_HO_Complete.response to Mobile Node.

3.1.2. Network-initiated Handover procedure

The Network-initiated Handover procedure operates as follows [2, 6]:
1) Serving PoS sends MIH_Get_Information.request to Information Server to get neighboring network information and Information Server responds by sending MIH_Get_Information.response.
2) Serving PoS triggers a network-initiated handover by sending MIH_Net_HO_Candidate_Query.request to Mobile Node. The MN responds through MIH_Net_HO_Candidate_Query.response which contains Mobile Node's acknowledgement about the handover and its preferred link and PoS lists.
3) Serving PoS sends MIH_N2N_HO_Query_Resources.request to one or more candidate PoSs to check the availability of the resource at candidate networks. Candidate PoS responds by sending MIH_N2N_HO_Query_Resources.response to Serving PoS.
4) Serving PoS decides the target of the handover based on the available resource status at candidate networks.
5) Serving PoS sends MIH_N2N_HO_Commit.request to target PoS to prepare resource at the target network. Target PoS responds the result of the resource preparation by sending MIH_N2N_HO_Commit.response.
6) After identifying that resource is successfully prepared, Serving PoS commands Mobile Node to commit handover toward the specified network type and PoA through MIH_Net_HO_Commit.request.
7) New layer 2 connection is established and Mobile Node sends MIH_Net_HO_Commit.response to Serving PoS.
8) After higher layer handover execution, Mobile Node may send MIH_MN_HO_Complete.request to Target PoS. Target PoS sends MIH_N2N_HO_Complete.request to previous Serving PoS to release resources which were allocated to Mobile Node. After identifying that resources are successfully released, target PoS may send MIH_MN_HO_Complete.response to Mobile Node.

3.2 USE CASE No. 1: HANDOVER FOR PROXY MIPv6

Figures 3.1 (a) and (b) show a network-initiated handover flow chart for Proxy Mobile IPv6 (PMIPv6), which is currently under standardization for supporting a local mobility in IETF NetLMM Working Group. The following handover flow refers to the overall flow and operates as follows:

1) MN receives packets through both Mobile Access Gateway (MAG) 1 located in the serving network and Local Mobility Anchor (LMA), which are primary components of the PMIPv6.

2) The Serving PoS queries the Information Server to get information about available neighboring networks.

3) The Serving PoS triggers a network-initiated handover by sending the MIH_Net_HO_Candidate_Query.request message to the MN. The MN responds through the MIH_Net_HO_Candidate_Query.response message, which contains MN's acknowledgement about the handover initiation and its preferred link and PoS lists.

4) The Serving PoS sends the MIH_N2N_HO_Query_Resource.request messages to different candidate PoSs (they can be more than one) to query the availability of the resource at candidate networks. The candidate PoSs respond by sending the MIH_N2N_HO_Query_Resource.response message to the Serving PoS. The Serving PoS decides the handover target based on the resource availability information of candidate networks informed by the MIH_N2N_HO_Query_Resource.response message.

5) The serving PoS informs the decided target PoS (i.e. Candidate Network 1 in the Figure 3.1(a), where MAG2 is located) of the handover commitment and requests the target PoS to prepare resources for the incoming MN through sending the MIH_N2N_HO_Commit.request message. The target PoS replies the result of the handover commitment and resource preparation by sending MIH_N2N_HO_Commit.response. (Upon receiving the MIH_N2N_HO_Commit Request message, PMIPv6 client in the target PoS may query the incoming MN's profile to an AAA server and send Proxy Binding Update in order to register the location of the MN in advance. The PMIPv6 client in the target PoS also may buffer the packets received from LMA until the MN attaches to the target PoS.)

6) The serving PoS requests MN to perform handover to the decided target PoS by sending the MIH_Net_HO_Commit.request message. The MN replies the result of the handover commitment by sending MIH_Net_HO_Commit.response message.

7) Upon detecting MN's detachment, PMIPv6 client in the Serving PoS terminates a current binding of the MN via sending Proxy Binding Update with Lifetime set to 0 and requests LMA to buffer packets destined for the MN.

8) Once the MN establishes Layer 2 connection to the target PoS, PMIPv6 client in the target PoS registers the current MN's location to LMA by sending Proxy Binding Update message. LMA updates its Binding Cache Entry with the Proxy Binding Update message and then replies with Proxy Binding Acknowledgement message. LMA also may forward the buffered packets.

9) After receiving the Proxy Binding Acknowledgement message, PMIPv6 clients send Router Advertisement message to the MN. The Router Advertisement is constructed with the MN's information obtained from the policy server and LMA. It can be solicited by Router Solicitation message from the MN or periodically transmitted. MN configures IP addresses on its interface, which is currently used to connect to the target PoS, with the received Router Advertisement message. Once the PMIPv6 procedures are completed, MN receives packets through both MAG 2 and LMA.

10) After the PMIPv6 execution, the target PoS sends the MIH_N2N_HO_Complete.request message to the previous serving PoS. The previous serving PoS responds the message with MIH_N2N_HO_Complete.response.

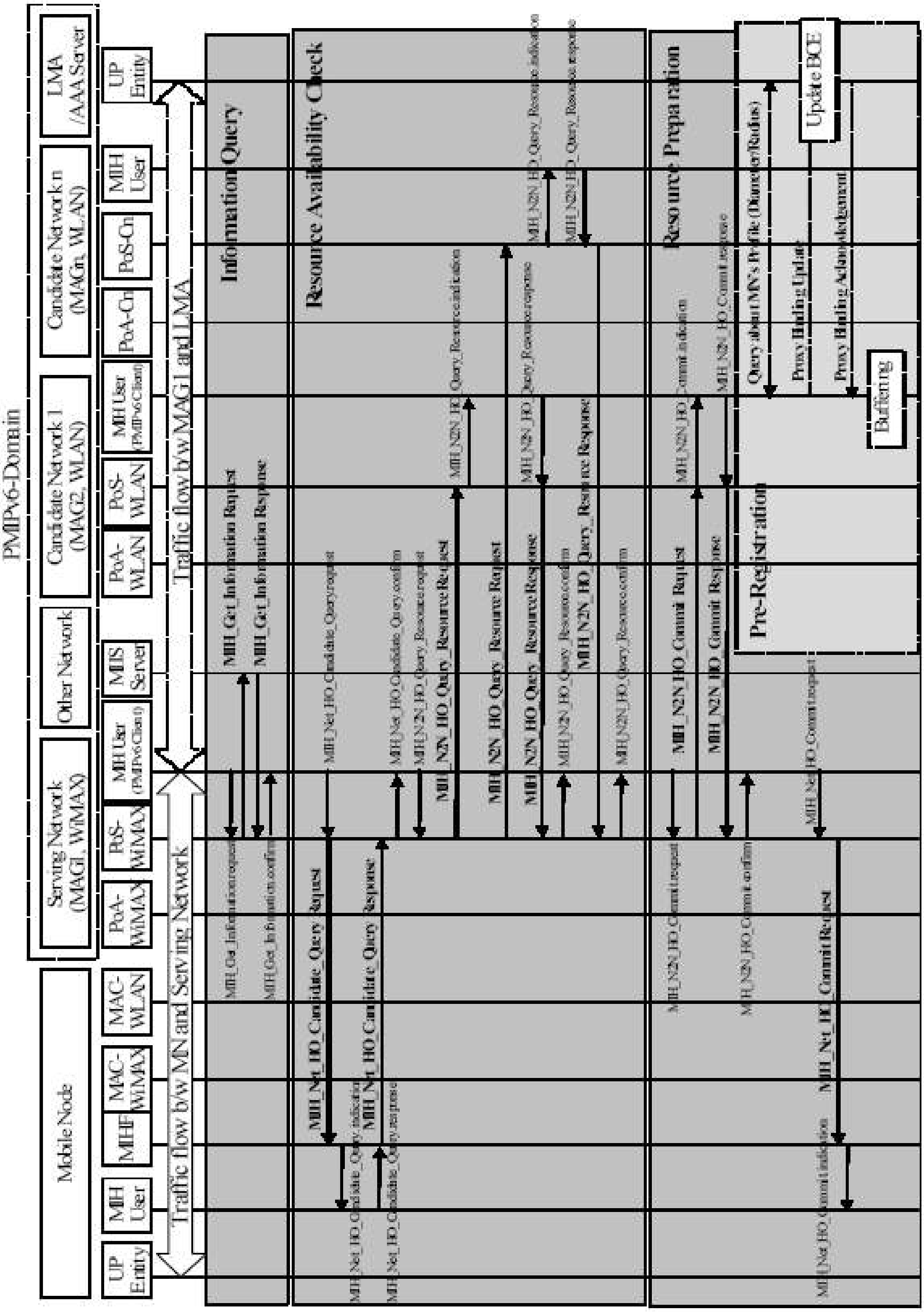


Figure 3.1 (a): Network-initiated handover procedure for Proxy MIPv6 [6]

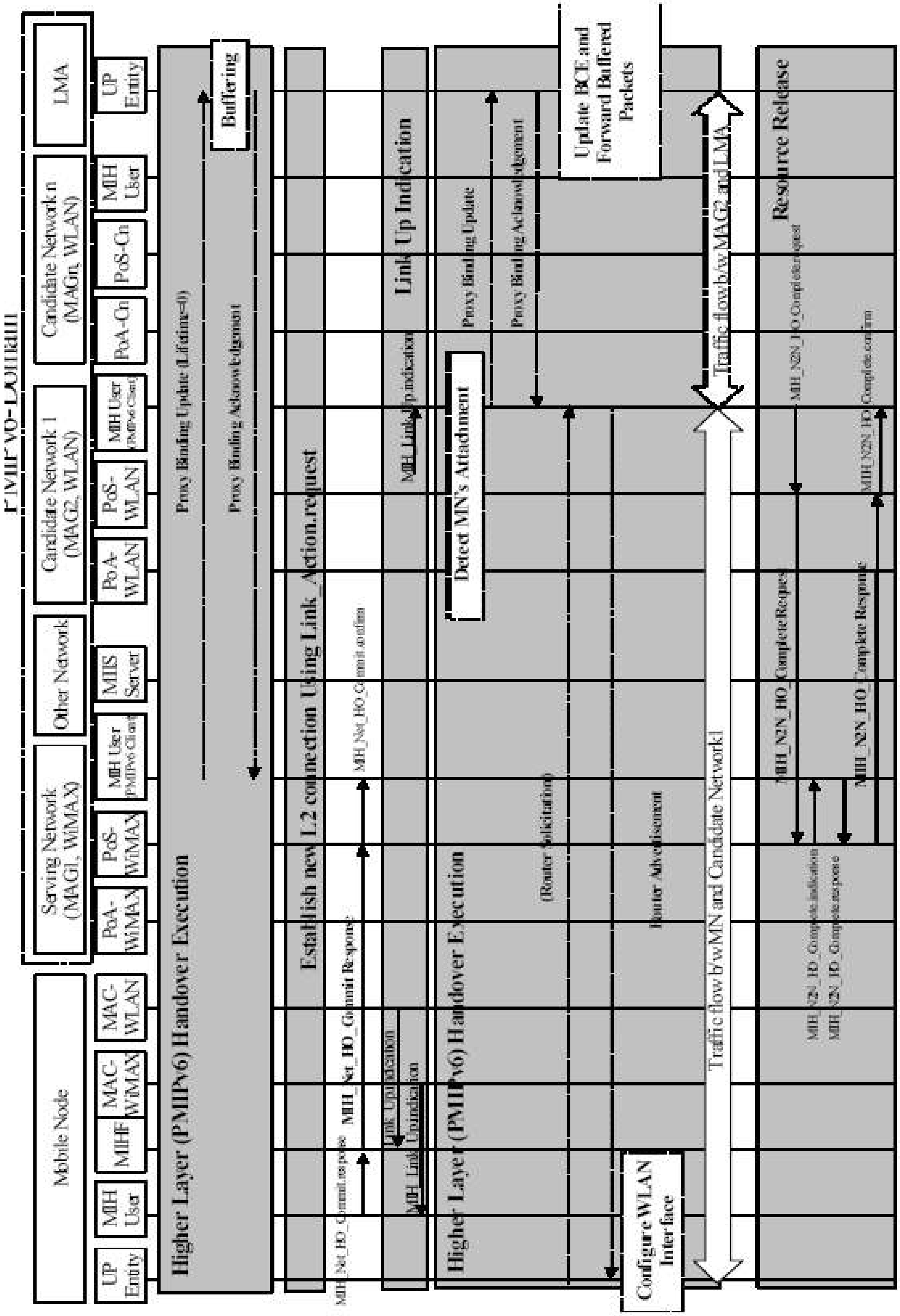


Figure 3.1 (b): Network-initiated handover procedure for Proxy MIPv6 (cont.)

3.3 USE CASE No.2: WiMAX/UMTS HANDOVER PROCEDURE

In this section, we try to present a specific scenario where a UE performs a vertical handover from WiMax to UMTS network. The procedure follows the general steps described in section 3.1. This section's contribution (as well as section 3.4 and 3.5) is their more coherent and specified presentation of the topics we have discussed in chapters 2 and 3 in order one can more circumstantially understand the overall handover procedure. In this scenario, we present all the three major steps of a handover.

According to the international literature, a handover in IEEE 802.21 consists of three major steps: initiation, preparation, and execution. In our example (Figure 3.2), the handover initiation phase involves the configuration of wireless devices to generate appropriate triggers towards the network and inform it about important changes in the link quality (step 1 in figure 3.2). This means that, at some point in time after the connection has been established, the WiMAX PoS issues a MIH_Configure_Link.request message towards the UE to define the required thresholds under which the WiMAX wireless device will report signal strength measurements. In fact, this message indicates the required QoS parameters for the link that are then conveyed to the wireless device and define the thresholds for which a report will be generated. The successful configuration at the wireless device is confirmed at the WiMAX PoS with a MIH_Configure_Link.confirm message.

When the previously defined thresholds are crossed, measurement reports (MIH_Link_Parameters_Report.indication) are sent to the WiMAX PoS. These messages can be sent periodically for informational reasons. However, when the deterioration in link quality is unacceptable (step 2), another type of message can be sent to indicate the necessity of the handover (MIH_Link_Going_Down.indication). This message indicates that the link will go down in a specific time interval and with a specific confidence. It also indicates the reason of this link unavailability.

After this indication, the handover preparation phase begins. This phase may include queries to the MIIS server for information retrieval related to the access networks close to the UE, additional scanning commands towards the UE, a preferences exchange between the UE and the WiMAX PoS, a resource availability check at the candidate networks, handover decision, and resource reservation. More specifically, immediately after receiving an indication for handover, the WiMAX PoS asks the MIIS server for information about available networks near the UE with a MIH_Get_Information.request message (step 3). This query is based on the current PoA location, and its result returns the candidate networks (PoAs) near the UE, along with their characteristics (MIH_Get_Information.response). The last information is quite useful because it can help the UE search for specific radio networks rather than every available network.

At this point, the UE is still unaware of the networks that are reachable in terms of RSS. Therefore, the WiMAX PoS orders it to begin scanning for PoAs mentioned in the response from the MIIS server, and the signal measurements are sent back to the WiMAX PoS (MIH_Scan.request and MIH_Scan.confirm messages, respectively). In addition, the WiMAX PoS issues a MIH_Net_HO_Candidate_Query.request message to verify the intention of the UE to hand over. With this message, the WiMAX PoS also proposes target networks to the UE and is informed about the UE preferences (MIH_Net_HO_Candidate_Query.response). This means that the preferred link list of the UE may differ from the suggested new link list of the network. In parallel, the network checks for available resources in the candidate networks (step 4). The WiMAX PoS communicates with its peer entity in the UMTS (UMTS PoS) for this reason (MIH_N2N_HO_Query_Resources.request and reply). It is important to note here that the translation of resource information across different technology networks is still an open and very interesting issue in IEEE 802.21.

Having all the information in the network and the UE side related to the handover, the WiMax PoS can decide about the target access network (step 5) and reserve resources (step 6). If resource reservation is guaranteed both in the new access network and the core network, the UE may enjoy a seamless or make-before-break handover. After the reservation stage, the handover execution phase starts, and the WiMAX PoS orders the UE to begin handover by indicating the actions over the old and the new link (MIH_Net_HO_Commit.request). Resources over the old link either can be maintained or released. This means that if both the old and new links are active concurrently, the handover interruption is minimized. In any case, the actions over the new link indicate to set up a new PDP context and perform mobile IP registration or a binding update between the UE and the home agent (HA) in the P-GW. The successful establishment of an IP connection over the UMTS is reported to the WiMAX PoS (MIH_Net_HO_Commit.response) and afterwards, to the UMTS PoS through a MIH_N2N_HO_Complete.request message. Moreover, resources over WiMAX are released if not released earlier (MIH_N2N_HO_Complete.response). After this, data can start flowing over the UMTS network.

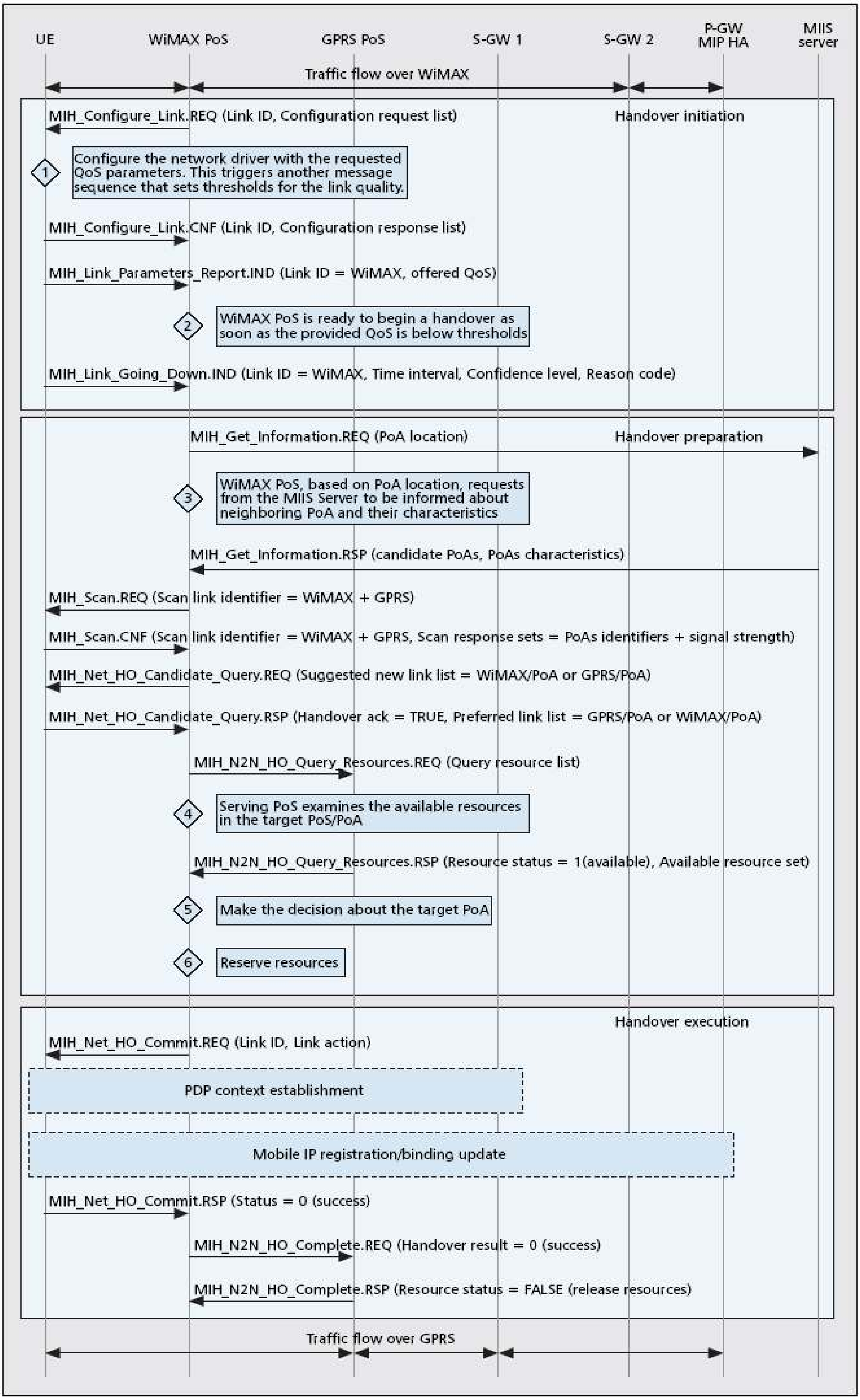


Figure 3.2: WiMax to UMTS handover procedure [3]

3.4 USE CASE No.3: UMTS/WLAN HANDOVER PROCEDURE

In this scenario, we selected the mobile-initiated case in order to present a UMTS/WLAN handover paradigm. Figure 3.3 shows the message exchange involved in a mobile-initiated handover from UMTS to WLAN. In the following a detailed explanation of the messages and procedures is presented.

The handover procedure starts by the MIH user of the mobile node (MN) querying the MIHF located on the MN itself about the surrounding networks (message 1 in figure 3.3). This query is forwarded by the MIHF to the information server located in the operator network (or a third party network). The query is started by message 1 and answered by message 4. Through these four messages the MN gets the required information in order to gain an understanding of the networks to which perform a handover while roaming through this specific geographical area. As the answer contains information regarding a possible WLAN network, the MN switches on its WLAN interface and starts listening for beacons.

Once a beacon is received, the IEEE 802.11 link layer will generate a Link_Detected.indication event (message 5). The link layer, through an IEEE 802.11 defined primitive, indicates the detection of a new link. This primitive is mapped into the event through the use of the MIH_LINK_SAP (see details in subsection 2.4.1). This indication is forwarded by the MIHF to the MIH user in message 6. When the MIH user receives the MIH_Link_detected.indication, it triggers the mobile-initiated handover by sending to its PoS (located on the 3G network) the information regarding potential candidate networks discovered up to the moment. This information is sent in message 7 to the MIHF, which forwards this query to the serving PoS (message 8).

After receiving message 8, the serving PoS starts querying the available candidate networks (taking into account the information provided by the MN) asking for the list of resources available and including the QoS requirements of the user (exchange 9). This is performed by a successive exchange of query resources messages with one or several candidate PoSs. The result of the queries is sent to the MN through message 10 and 11. At this point, the MN has enough information about the surrounding networks to take a decision on the network to which hand over. Once the MIH User has decided the target network to hand over, it delivers a switch command to the MIHF (message 12), which will trigger a WLAN layer 2 (L2) connection. After issuing the commands to start WLAN connection establishment, the MIHF sends an event to the MIH user indicating the start of the connection (message 13). Once the connection is established, the WLAN MAC layer issues an event reporting the end of the L2 handover to the MIHF (message 14), which will be forwarded to the MIH user (message 15).

Once message 15 is received, a higher-layer handover procedure can start. In this case Mobile IP has been selected (see more details in section 3.2), although any other mobility management protocol would be equally suited. When the handover is completed at the higher layers, the MIH user sends a MIH_HO_Complete message to the MIHF, which will inform the target PoS (messages 16 and 17), which becomes the new serving PoS. At this point the target PoS informs all the implied NEs of the handover finalization (exchange 18). Specifically, the target PoS has to inform the serving PoS of the handover completion so it can release any resources. Finally, message 19 closes the handover procedure, indicating to the MN that the procedure has finished, and message 20 informs the MIH user.

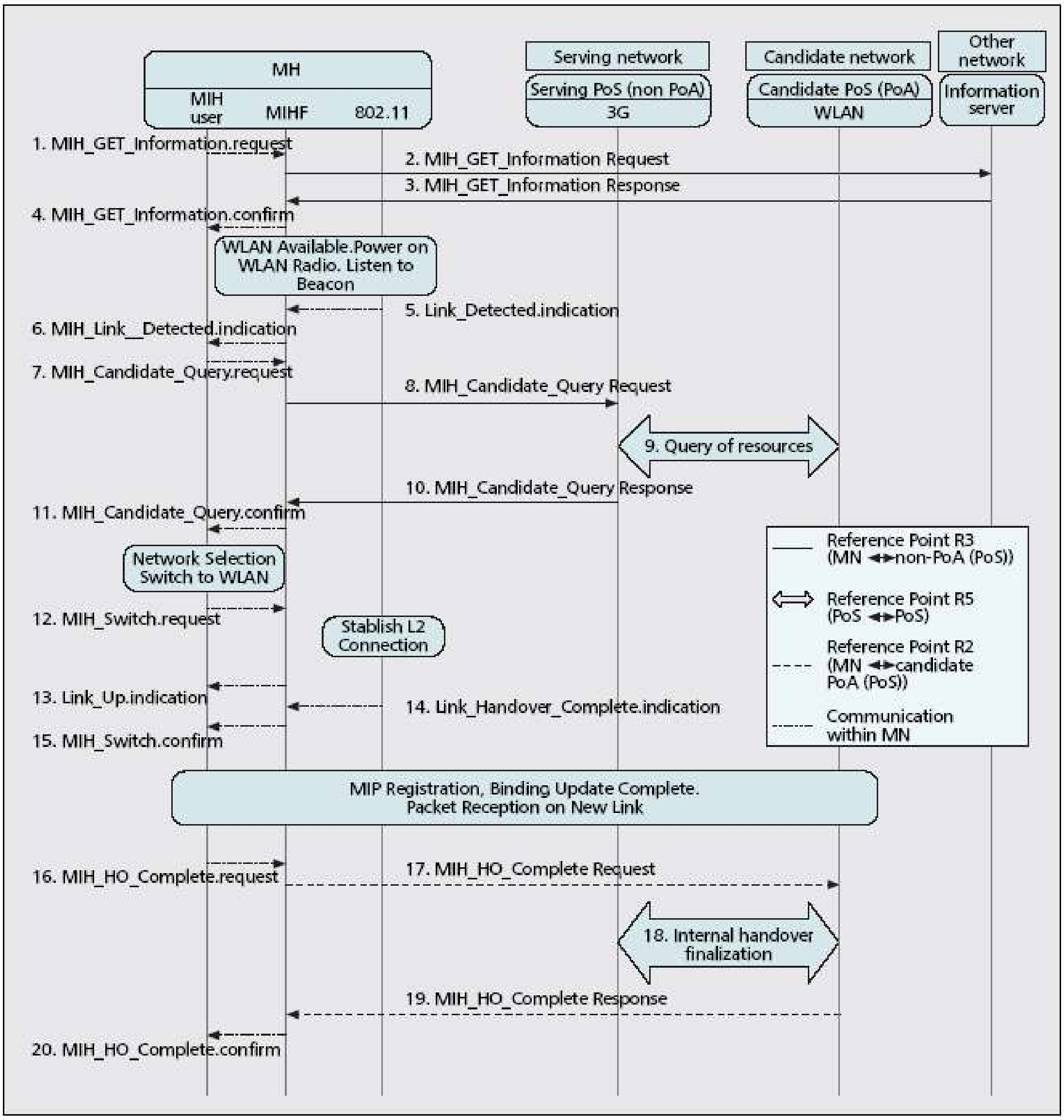


Figure 3.3: UMTS to WLAN handover procedure [4]

3.5 USE CASE No.4: WiMAX/WLAN HANDOVER PROCEDURE

In this scenario, we assume a WiMAX/WLAN interworking paradigm. We also assume that a mobile-initiated handover procedure takes place (see more details about the general principles of mobile-initiated handover in subsection 3.1.1). Figure 3.4 provides an example call flow or messaging diagram to illustrate how IEEE 802.21 media-independent handover (MIH) services can facilitate the handover in the case of transitioning from a high-mobility low data rate network to a low-mobility high data rate one.

For this example we first assume the mobile-controlled mobility model, that is, the mobility decision is made by an IEEE 802.21-enabled mobile node (MN). We further assume that the MN is initially travelling at high velocity and is connected to the IEEE 802.16m access network (see step 1 in figure 3.4). After that, the mobility management decision logic on the MN decides that it needs to prepare for a possible handover to an IEEE 802.11VHT network (step 2). It may realize that its velocity is decreasing. Or the user or one of the applications running wants to explore the possibility of getting higher data throughput from an 802.11 VHT network.

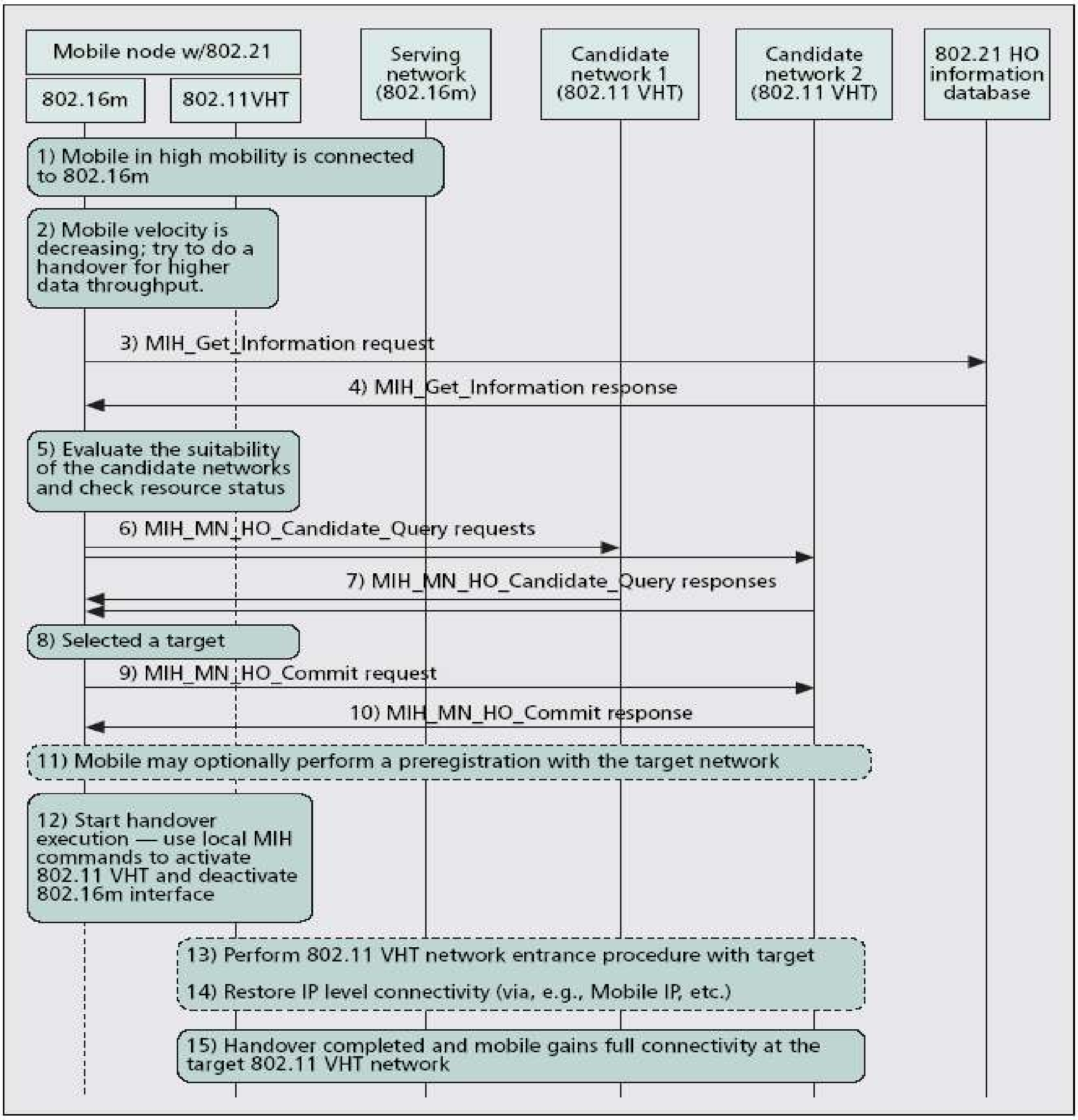


Figure 3.4: WiMAX to WLAN handover procedure [5]

The MN then uses IEEE 802.21 information service primitives (steps 3 and 4) to query a network-based IEEE 802.21 handover information database. It does this to both identify and get information on IEEE 802.11 VHT network(s) with coverage in its vicinity. Example information that may be available on these networks is identity/owner, service agreement, base station locations, QoS grades, and data rate (see more details in subsection 2.4.2.8). We assume in the following that there are two IEEE 802.11 VHT networks available in the MN's vicinity. They are the candidates for the handover target network. After obtaining information on the two candidate networks, the MN (step 5) uses its own criteria to evaluate suitability of each as the handover target network. For example, a candidate could be eliminated for lack of an adequate service agreement.

In addition, the MN may also decide to check the resource status of each candidate before making a choice. To do so, the MN (steps 6 and 7) uses IEEE 802.21 MIH command service messages or primitives (command services are further described in subsection 2.3.2) to query each candidate. In the query the MN indicates to each candidate the minimal set of resources necessary for sustaining its current application sessions after a handover. Taking into account resource status at each candidate, the MN (step 8) decides which candidate will be the handover target network. For this example, we assume that the MN selects candidate network #2. The MN (steps 9 and 10) then notifies the selected target network with IEEE 802.21 MIH_MN_HO_Commit primitives, requesting the reservation of the needed handover resources at the target network.

At this point, the MN (step 11) may also perform some preregistration procedures (e.g., pre-authentication, pre-allocation of IP address, etc.) at the target network, in order to reduce the overall handover latency. At step 12, the MN decides to start execution of the handover to the target, and invokes IEEE 802.21 commands to activate its IEEE 802.11 VHT interface and deactivate its IEEE 802.16m interface. After this, the handover enters the execution phase, which is not particularly in the scope of 802.21 standard.

3.6 REFERENCES

[1] Joo-Young Baek et al, "**Network-initiated Handover Based on IEEE 802.21 Framework for QoS Service Continuity in UMTS/802.16e Networks**", Vehicular Technology Conference, May 2008.

[2] IEEE P802.21/D9.0, "**Draft Standard for Local and Metropolitan Area Networks: Media Independent Handover Services**", February 2008.

[3] George Lampropoulos, Apostolis Salkintzis and Nikos Passas, "**Media-Independent Handover for Seamless Service Provision in Heterogeneous Networks**", IEEE Communications Magazine, January 2008.

[4] Antonio de la Oliva et al, "**An Overview of IEEE 802.21: Media-Independent Handover Services**", IEEE Communications Magazine, August 2008.

[5] Les Eastwood, Scott Migaldi, Qiaobing Xie and Vivek Gupta, "**Mobility Using IEEE 802.21 in a Heterogeneous IEEE 802.16/802.11-Based, IMT-Advanced (4G) Network**", IEEE Wireless Communications, April 2008.

[6] https://mentor.ieee.org/802.21/file/07/21-07-0161-03-0000-pmip-flowchart.doc

CHAPTER 4
IEEE P1900.4: RECONFIGURATION OF MULTI-RADIO SYSTEMS

4.1 SCOPE – PURPOSE – NEED FOR THE STANDARD

The scope of the standard is to define the building blocks comprising a) network resource managers, b) device resource managers, and c) the information to be exchanged between the building blocks, for enabling coordinated network-device distributed decision making which will aid in the optimization of radio resource usage, including spectrum access control, in heterogeneous wireless access networks. The standard is limited to the architectural and functional definitions at a first stage. The corresponding protocols definition related to the information exchange will be addressed at a later stage.

The purpose is to improve overall composite capacity and quality of service of wireless systems in a multiple Radio Access Technologies (RATs) environment, by defining an appropriate system architecture and protocols which will facilitate the optimization of radio resource usage, in particular, by exploiting information exchanged between network and mobile terminals, whether or not they support multiple simultaneous links and dynamic spectrum access.

Multimode reconfigurable devices are increasingly being adopted within the wireless industry. The choice among various supported air interfaces on a single wireless device is already a reality today, with devices offering, for example, 2nd, 3rd generation cellular radio access technologies and 802 wireless standards. Further, devices and networks with dynamic spectrum access capabilities allowing the use of spectrum resource simultaneously among different systems are emerging (e.g. IEEE 802.22) and will be part of the radio eco-space. There is a need to define a standard addressing the overall system architecture and information exchange between the network and devices, which will allow devices (access points and end-user terminals) to optimally choose among the available radio resources and simultaneously use several of these resources such that the overall efficiency and capacity of the resulting composite network is improved [1, 2].

4.2 OVERALL SYSTEM DESCRIPTION

Terminal equipment in radio access networks (RANs) are becoming increasingly intelligent and spectrum aware. The ability to perform dynamic spectrum access at the terminal level and manage spectrum at the RAN level is moving RAN to more intelligent and cognitive operation. The DSASM enables dynamic spectrum resource usage utilization. The DSASM standard addresses the problem of RAN interoperability in heterogeneous wireless environments that include multiple RANs that are operated by different operating companies. This section describes the overall description of the RAN environment and the operations required to manage dynamic spectrum access in the heterogeneous wireless environment.

The field of application of P1900.4 is radio systems with: a) Composite wireless network environment (multiple Radio Access Networks and multiple radio interfaces), b) Cognitive terminals with multi-mode and multi-link capabilities, c) Multiple operators or meta-operator regrouping several operators. Within this field of application, the standard provides common means to improve overall composite capacity and quality of service, facilitate optimization of radio resource usage, support reconfiguration capabilities of RANs and terminals, collect RAN-related and terminal-related context information, support information exchanged between network and terminals for distributed decision making and request reconfiguration of RANs and terminals [2].

The IEEE P1900.4 system will comprise a logical channel through which the network and terminals communicate via their respective Reconfiguration Management Entities (RMEs). The RME on the network side is termed the Network Reconfiguration Manager (NRM), and the RME on the terminal side is the Terminal Reconfiguration Manager (TRM). This logical channel is called Radio Enabler (RE) [3]. The Radio Enabler may be mapped onto: a) one or several legacy RANs, already used for data transmission, or b) one or several dedicated RANs. While in the former case the RE can be materialized from existing channels, e.g., Broadcast Control Channels (BCCH), as defined in the GERAN (in-band solution); in the latter case, the Radio Enabler is mapped onto a newly defined, dedicated physical channel and equipment, as an out-of-band channel. In the following figure (figure 4.1), we can see an example of a composite wireless network within which the P1900.4 system is deployed [2].

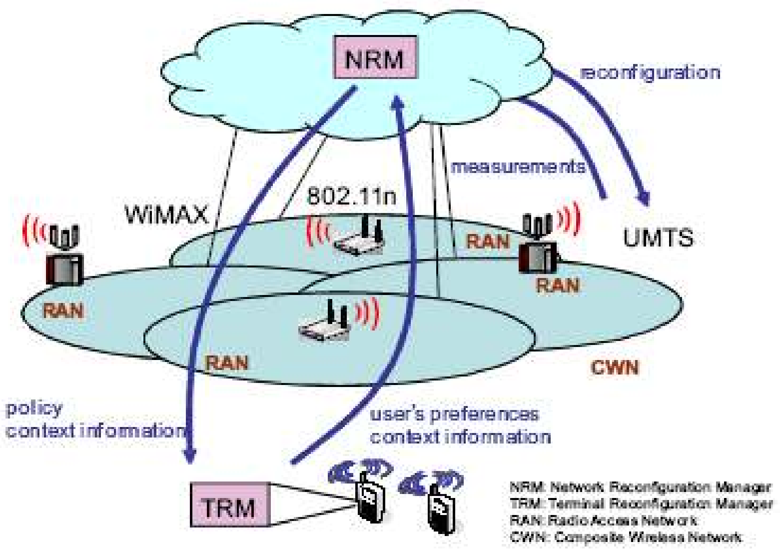


Figure 4.1: The concept of IEEE P1900.4 standard [4]

IEEE P1900.4 considers a heterogeneous wireless environment involving multiple RANs (which may be of different RATs) belonging to different network operators. The considered environment may also comprise the case of a “composite network”, consisting of multiple RANs (which again may be of different RATs), under the control of a single network operator. Cognitive Terminals (CTs), having multi-mode and/or multi-homing capability, facilitate the standard through being able to connect to one or several RANs/RATs of one or several operators. There are, however, foreseeable scenarios in which the standard remains applicable even if CTs are not present.

In this section, we simply try to summarize the reference use cases which P1900.4 is designed to serve. Chapter 5 provides more details about the three uses cases defined within IEEE P1900.4, that is a) Dynamic Spectrum Assignment, b) Dynamic Spectrum Access and c) Distributed Radio Resource Usage Optimization.

In the Dynamic Spectrum Assignment use case, frequency bands are dynamically assigned to the RANs among the participating networks in order to optimize spectrum usage. In other words, the assigned frequency bands are not fixed, and can be dynamically changed. OSMs generate spectrum assignment policies expressing regulatory framework and operators intentions regarding possible changes to the spectrum assignment. OSMs provide these spectrum assignment policies to corresponding NRMs. NRMs analyze spectrum assignment policies and available context information and dynamically make spectrum assignment decisions to improve spectrum usage and quality of service. After the new spectrum assignment decisions, NRMs request corresponding reconfiguration of their RANs. Following RANs and BSs reconfiguration, cognitive terminals need to reconfigure correspondingly. Example manifestations for this Dynamic Spectrum Assignment use case are spectrum sharing, and spectrum renting. In the spectrum renting example, frequency bands of one RAN are assigned to another RAN on a temporary basis. In the spectrum sharing example, one frequency band is shared among several RANs in accordance with the dynamic spectrum assignments. Single-operator and multi-operator scenarios are possible within the Dynamic Spectrum Assignment use case. Within the multi-operator scenario, the NRM is considered as being located either inside or outside of the operator (see more details in chapter 5).

In the Dynamic Spectrum Access use case, frequency bands assigned to RANs are fixed. However, a particular frequency band can be shared by several RANs. In other words, the Dynamic Spectrum Access use case describes how fixed frequency bands are shared and/or used dynamically by RANs and cognitive terminals. NRMs analyze available context information and dynamically make spectrum access decisions to improve spectrum usage and quality of service. NRMs make these spectrum access decisions within the framework defined by spectrum assignment policies. Following these decisions, NRMs request corresponding reconfiguration of their RANs. NRMs dynamically generate radio resource selection policies and send them to their TRMs. These radio resource selection policies will guide these TRMs in their spectrum access decisions. TRMs analyze radio resource selection policies and available context information and dynamically make spectrum access decisions to improve spectrum usage and quality of service. TRMs make these spectrum access decisions within the framework defined by radio resource selection policies. Following these decisions, TRMs request corresponding reconfiguration of their cognitive terminals.

Distributed Radio Resource Usage Optimization use case demonstrates how the IEEE P1900.4 system can be applied to legacy RANs in order to optimize radio resource usage and improve quality of service. In the Distributed Radio Resource Usage Optimization use case, frequency bands assigned to RANs are fixed. Also, BSs are not capable to reconfigure. Distributed Radio Resource Usage Optimization use case considers cognitive terminals with multi-mode and/or multi-homing capability. Multi-mode cognitive terminals can have only one active connection. However, they can dynamically reconfigure this active connection accessing different RANs with different radio interfaces. Multi-homing terminals can have multiple simultaneous active connections with different RANs. Also, they can dynamically reconfigure these connections. Decision on cognitive terminal reconfiguration is made by its TRM and is supported by NRM. NRMs analyze available context information, dynamically generate radio resource selection policies, and send them to their TRMs. These radio resource selection policies will guide these TRMs in their spectrum access decisions. TRMs analyze radio resource selection policies and available context information and dynamically make decisions on reconfiguration of their cognitive terminals to improve spectrum usage and quality of service. TRMs make these reconfiguration decisions within the framework defined by radio resource selection policies. Following these decisions, TRMs request corresponding reconfiguration of their cognitive terminals.

4.3 REQUIREMENTS

There are three key features in P1900.4 as we describe below. The first feature is that the network side such as operators or service providers does not make any decision on terminal reconfiguration. The network side does not transmit policy information which terminals need to obey. Instead of this, terminals make decisions by themselves on their reconfigurations. The network side transmits just supporting information for terminals as recommendations. The terminals make decisions taking into account the recommendation as well as user's preferences and applications' requirements. Users need to find and adopt good algorithms for this decision making and information for them. Network servers that create the recommendation can be deployed even by users.

The second feature is multi link aggregation crossing operators dynamically adapting to radio link QoS according to applications in use. As QoS parameters to be considered depend on applications, terminals detect application type from traffic patterns, and select the best combination of radio links. There are two types of link aggregation; distribution and duplication. For example, in case of voice communication, it would be better to duplicate packets rather than distribute because traffic bandwidth is very small and jitters should be avoided. On the other hand, in case of file downloading, traffic distribution onto different links would be better.

The third feature is end-to-end QoS support. Taking into consideration end-to-end communication quality, terminals decide and execute their reconfigurations. For this purpose, networks measure end-to-end QoS and provide the results via a common interface. In case of P1900.4, RAN related measurement is considered, but quality in data communication network between RANs is not [4].

4.3.1 System Requirements

4.3.1.1 Context Awareness

There shall be entities on network side and terminal side responsible for context information collection. Context information collection entity on network side shall collect RAN-related context information. RAN-related context information may include: a) RAN radio resource optimization objectives, b) RAN capabilities, c) RAN-related measurements, d) RAN transport capabilities and e) other RAN-related context information. NRM shall be able to obtain RAN-related context information from context information collection entity on network side. NRM may receive this context information periodically and/or in response to request from NRM and/or on event. Context information collection entity on network side may be implemented in a distributed manner.

Context information collection entity on terminal side shall collect terminal-related context information. Terminal-related context information may include: a) user's preferences, b) required QoS levels, c) terminal's capabilities, d) terminal-related measurements, e) geo-location information, f) geo-location based terminal-related measurements and g) other terminal-related context information. TRM shall be able to obtain terminal-related context information from context information collection entity on terminal side.

NRM and TRM shall exchange context information. NRM shall send the following context information to TRM: a) RAN-related context information and b) terminal-related context information related to other terminals. NRM may send this context information to TRM periodically and/or in response to request from NRM and/or on event. TRM sends the terminal-related context information to NRM. TRM may send this context information to NRM periodically and/or in response to request from NRM and/or on event.

4.3.1.2 Decision Making

Decision making in this standard is based on policy-based management framework. There shall be entity on network side responsible for generating spectrum assignment policies. NRM shall be able to obtain spectrum assignment policies from this entity. Spectrum assignment policies adhere to regulatory framework. Spectrum assignment policies may express operators' needs and possibilities related to DS Assignment. NRM makes DS Assignment decisions compliant with received spectrum assignment policies. The NRM provides information on its DS Assignment decisions to the legacy entity on network side within the operator's control, which is responsible for generating spectrum assignment policies. NRM makes DS Access decisions compliant with its DS Assignment decisions and with received spectrum assignment policies. NRM generates radio resource selection policies and provides radio resource selection policies to TRM.

Radio resource selection policies and context information should be sent to an optimized set of TRMs in order not to overload selected RANs with broadcast. TRM makes decisions on terminal reconfiguration compliant with received radio resource selection policies. If specified by radio resource selection policies, TRM of geo-localization-capable terminal makes decisions based on terminal geo-location. NRM may have means to specify geo-location based radio resource selection policies to TRM.

4.3.1.3 Reconfiguration

There shall be entities on network side and terminal side responsible for reconfiguration. NRM sends reconfiguration requests to reconfiguration entity on network side. Following received reconfiguration requests, reconfiguration entity on network side requests and controls reconfiguration of RANs. Reconfiguration entity on network may be implemented in a distributed manner. TRM sends reconfiguration requests to reconfiguration entity on terminal side. Following received reconfiguration requests, reconfiguration entity on terminal side requests and controls reconfiguration of terminal.

4.3.2 Functional requirements

4.3.2.1 NRM Functionality

NRM shall have capability to make DS Assignment and DS Access decisions. NRM shall also have capability to request reconfiguration of RANs and BSs corresponding to these DS Assignment and DS Access decisions. NRM shall have capability to evaluate the efficiency of spectrum usage under the current spectrum assignment. The evaluation results shall be made available to assist in improving the efficiency of future DS Assignment and DS Access decisions. NRM shall have capability to generate radio resource selection policies. Radio resource selection policies should not contradict with each other. Radio resource selection policies shall be defined in a way that they correspond to a specific group of terminals (could be composed of any number of terminals). The targeted group of terminals should be defined based on the needs of CWN (Composite Wireless Network) radio resource usage optimization objectives, terminal location and radio resource usage patterns of the terminals obtained from the information collection, extraction and storage.

Radio resource selection policies shall guide TRMs in terminals reconfiguration decisions. NRM may specify policies referring to specific geo-location based terminal-related measurements and shall have the capability to setup and control the time interval for terminal reconfiguration. TRM shall perform terminal reconfiguration within this time interval. NRM shall have capability to evaluate the efficiency of current radio resource selection policies. The evaluation results shall be made available to assist in improving the efficiency of future radio resource selection policies. NRM shall have the capability to use the context information. That is, a) the RAN-related context information and b) the terminal-related context information. NRM shall have capability to select RANs for exchanging radio resource selection policies and context information between NRM and TRM.

Finally, NRM should have capability to make DS Assignment and DS Access decisions, request reconfiguration of RANs and BSs, evaluate the efficiency of spectrum usage, generate radio resource selection policies, evaluate the efficiency of current radio resource selection policies, and select RANs for exchanging radio resource selection policies and context information between NRM and TRM in a cooperative manner.

4.3.2.2 TRM Functionality

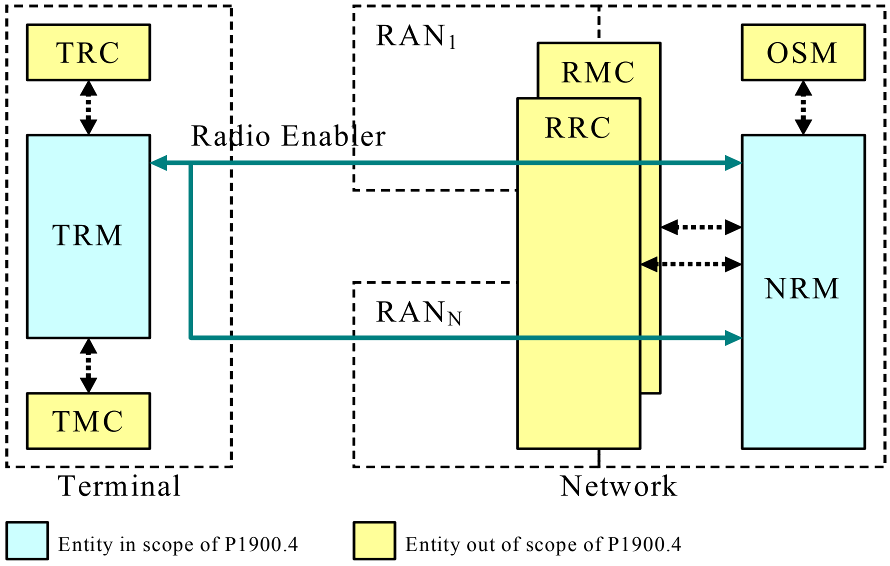
TRM shall have capability to make DS Access decisions and to request reconfiguration of its host terminal corresponding to these DS Access decisions. TRM shall also have capability to receive, process, and store the following context information: a) terminal-related context information, b) RAN-related context information.

TRM shall have capability to select RANs for exchanging radio resource selection policies and context information between NRM and TRM. NRM should have capability to make DS Access decisions, request reconfiguration of its terminal, and select RANs for exchanging radio resource selection policies and context information between NRM and TRM in a cooperative manner.

4.4 SYSTEM ARCHITECTURE

4.4.1 System Description

According to general system requirements described in the previous sections, the following system architecture is considered in this standard (figure 4.2).



TRM – Terminal Reconfiguration Manager
TRC – Terminal Reconfiguration Controller
RAN – Radio Access Network
RRC – RAN Reconfiguration Controller
NRM – Network Reconfiguration Manager
TMC – Terminal Measurement Collector
OSM – Operator Spectrum Manager
RMC – RAN Measurement Collector

Figure 4.2: System Architecture [4]

In Figure 4.2, the following key entities are identified: 1) the Terminal Reconfiguration Manager (TRM), 2) the Network Reconfiguration Manager (NRM), 3) the Terminal Reconfiguration Controller (TRC), 4) the Terminal Measurement Collector (TMC), 5) the RAN Reconfiguration Controller (RRC), 6) the RAN Measurement Collector (RMC), 7) the Operator Spectrum Manager (OSM), and 8) the Radio Enabler (RE). In this section we try to describe the general functionalities of all the entities mentioned above as well as the interfaces between these entities.

Terminal Reconfiguration Manager is the entity that manages the terminal for network-terminal-distributed optimization of spectrum usage within the framework defined by the NRM and in a manner consistent with users' preferences and available context information. Network Reconfiguration Manager is the entity that manages the Composite Wireless Network and terminals for network-terminal-distributed optimization of spectrum usage. NRM may be implemented in a distributed manner as it has been already explained in the previous section of this chapter.

Terminal Reconfiguration Controller is the entity that enables TRM reconfiguration of terminal. Terminal Measurement Collector is the entity that collects terminal-related context information and provides it to TRM. RAN Reconfiguration Controller is the entity that enables NRM reconfiguration of RANs. RAN Reconfiguration Controller may be implemented in a distributed manner. RAN Measurement Collector is the entity that collects RAN-related context information and provides it to NRM. It may be also implemented in a distributed manner. Operator Spectrum Manager (one entity per operator) is the entity that enables operator to control NRM DS Assignment decisions.

Radio Enabler (RE) is the logical communication channel between NRM and TRM. Radio Enabler may be mapped onto one or several RANs used for data transmission (in-band channel) and/or onto one or several dedicated RANs (out-of-band channel). All devices can listen to and likely transmit upon this radio enabler in order to obtain/convey reconfiguration management and associated information in as efficient a manner as possible. Among the many uses for this envisaged entity, devices might receive context information (including, for example, information about connectivity options upon start-up) far more efficiently, just by listening to a single channel instead of having to scan among multiple network pilots. As other benefits, reconfiguration processes can be coordinated among affected elements, the transferral of policies/directives and decisions related to a reconfiguration can be expedited or indeed made possible, and dynamic spectrum allocation technologies ensuing from reconfiguration, such as cognitive radio, can be facilitated through knowledge transferral (e.g. about spectrum usage) among heterogeneous devices and network elements. To summarize, the research community sees the radio enabler entity as being an essential component for future reconfigurable devices and network elements, and as being an important facilitator for dynamic spectrum access procedures both through assisting reconfiguration processes and through benefiting associated technologies such as cognitive radio [3].

As one can see from figure 4.2, there are many interfaces between the entities. The interface between the main entities, that is NRM and TRM, is used to transmit terminal-related context information from the TRM to NRM perspective. From NRM to TRM, transmissions of radio resource selection policies, RAN-related context information and terminal-related context information are taking place. In section 4.3, there is a thorough analysis of the context information just mentioned.

TRM, TRC and TMC are three entities lying at the terminal side. TRM transmits terminal-related context information requests to TRC and the latter responds to these requests with corresponding responses. TRM also transmits terminal-related context information requests to TMC. TMC provides TRM with terminal-related context information.

At the network side, NRM has several interfaces with entities such as the RRC, the RMC and the OSM. More specifically, it transmits RAN reconfiguration requests to RRC and receives corresponding responses. It also transmits RAN-related context information to RMC and receives corresponding responses. Finally, OSM transmits high-level spectrum assignment policies to NRM and the latter responds with information on spectrum assignment decisions being made by the Network Reconfiguration Decision and Control module of NRM entity (see more details in subsection 4.4.2).

4.4.2 Functional Description

According to the functional requirements, presented in subsection 4.3.2 of this chapter, the functional architecture which is considered in P1900.4 standard is demonstrated in figure 4.3.

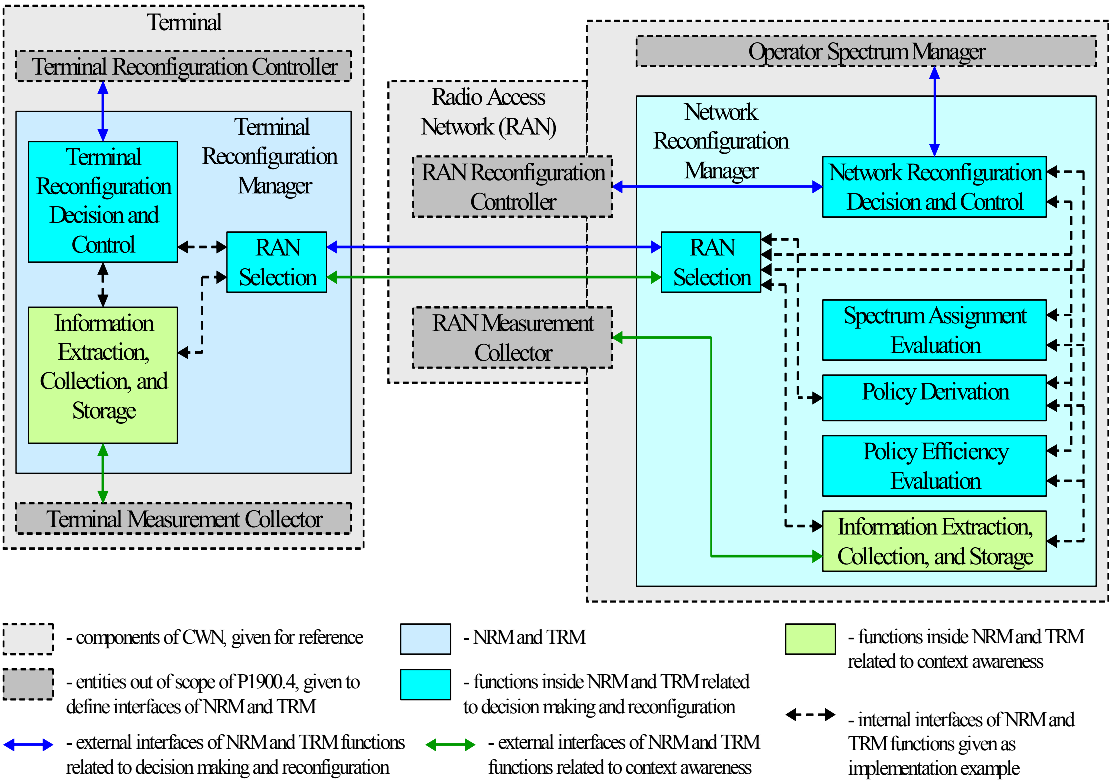


Figure 4.3: Functional Architecture of P1900.4 standard

4.4.2.1 NRM Functions and Interfaces

The following six main functions are identified inside NRM as one can see in figure 4.3: a) Policy Derivation, b) Policy Efficiency Evaluation, c) Network Reconfiguration Decision and Control, d) Spectrum Assignment Evaluation, e) Information Extraction, Collection, and Storage, and f) RAN Selection. In this subsection, we try to describe the main functionalities of these modules as well as to describe the way they cooperate with each other.

Policy Derivation function generates radio resource selection policies which guide TRMs in terminals reconfiguration decisions. The radio resource selection policies are derived according to context information from information collection, extraction and storage module. Policy Efficiency Evaluation function evaluates the efficiency of current radio resource selection policies. Evaluation results may be used by Policy Derivation function during generating radio resource selection policies.

Network Reconfiguration Decision and Control function makes decisions on RANs and BSs reconfiguration compliant with spectrum assignment policies received from OSMs. After making these decisions, Network Reconfiguration Decision and Control function sends corresponding reconfiguration commands to RRC. Also, Network Reconfiguration Decision and Control function sends information on made decisions to OSMs. Spectrum Assignment Evaluation function evaluates the efficiency of spectrum usage under the current spectrum assignment. Evaluation results may be used by Network Reconfiguration Decision and Control function during making decisions on RANs and BSs reconfiguration.

Information Extraction, Collection and Storage function receives, processes, and stores the RAN-related and the terminal-related context information. RAN-related context information is received from RMC. The information that is being included in the RAN-related context information has been already analyzed in subsection 4.3.1.1 of the current chapter. Terminal-related context information is received from TRM and for more details the reader should go back to the context awareness subsection of this chapter. Information Extraction, Collection and Storage function provides information to functions inside NRM and forwards RAN-related and terminal-related context information to TRM.

RAN Selection function selects RANs for exchanging radio resource selection policies and context information between NRM and TRM. Radio resource selection policies are sent from NRM to TRM. From NRM to TRM, RAN-related and terminal-related context information are sent. NRM receives terminal-related context information from TRM.

Policy Derivation, Policy Efficiency Evaluation, RAN Selection, Network Reconfiguration Decision and Control, and Spectrum Assignment Evaluation functions cooperate during their operation. They represent different aspects of decision making and reconfiguration. During their operation, these functions use information from Information Extraction, Collection, and Storage function. More specifically, the interfaces between NRM and other entities on network side which are identified in P1900.4 draft standard are: a) between Network Reconfiguration Decision and Control function and OSM, b) between Network Reconfiguration

Decision and Control function and RRC, and c) between Information Extraction, Collection and Storage function and RMC. The first interface is used in order OSM to send spectrum assignment policies and to receive NRM spectrum assignment decisions from Network Reconfiguration Decision and Control function. Interface (b) is used to transmit commands for RANs and BSs reconfiguration from Network Reconfiguration Decision and Control function to RRC. Eventually, the third interface is used to transmit RAN-related context information from RMC to Information, Extraction, Collection and Storage function.

4.4.2.2 TRM Functions and Interfaces

The following functions are identified inside TRM as it can also be seen in figure 4.3): a) Terminal Reconfiguration Decision and Control, b) Information Extraction, Collection, and Storage, and c) RAN Selection.

Terminal Reconfiguration Decision and Control function makes decisions on terminal reconfiguration. After making these decisions, it sends corresponding reconfiguration commands to TRC. Information Extraction, Collection and Storage function receives, processes, and stores the terminal-related and the RAN-related context information (see more details in subsection 4.3.1.1). Terminal-related context information is received from TMC. Terminal-related context information regarding other terminals is received from NRM. On the other hand, RAN-related context information is received from NRM.

Information Extraction, Collection and Storage function provides information to functions inside TRM. Information Extraction, Collection and Storage function forwards terminal-related context information to NRM. RAN Selection function selects RANs for exchanging radio resource selection policies and context information between NRM and TRM. Radio resource selection policies are sent from NRM to TRM. From NRM to TRM, RAN-related context information and terminal-related context information are sent. From TRM and NRM, terminal-related context information is sent.

Terminal Reconfiguration Decision and Control and RAN Selection functions cooperate during their operation. They represent different aspects of decision making and reconfiguration. During their operation, these functions use information from Information Extraction, Collection and Storage function. Terminal Reconfiguration Decision and Control function makes reconfiguration decisions within the framework determined by the received radio resource selection policies. More specifically, there are two main interfaces defined between TRM and other entities on terminal side: a) between Information Extraction, Collection and Storage function and TMC, and b) between Terminal Reconfiguration Decision and Control function and TRC. The first interface is used to transmit terminal-related context information from TMC. The second interface is used to transmit commands for terminal reconfiguration from Terminal Reconfiguration Decision and Control function to TRC.

4.5 REFERENCES

[1] IEEE P1900.4/D1.5, "**Draft Standard for Architectural building blocks enabling network-device distributed decision making for optimized radio resource usage in heterogeneous wireless access networks**", April 2008.

[2] Soodesh Buljore, Hiroshi Harada et al, "**IEEE P1900.4 Standard: Reconfiguration of Multi-Radio Systems**", Computational Technologies in Electrical and Electronics Engineering conference, Sibircon, 21-25 July 2008.

[3] Oliver Holland, Markus Muck et al, "**Development of a Radio Enabler for Reconfiguration Management within the IEEE P1900.4 Working Group**", 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN, 17-20 April 2007.

[4] Kentaro Ishizu, Homare Murakami et al, "**Design and Implementation of Cognitive Wireless Network based on IEEE P1900.4**", Sensor, Mesh and Ad Hoc Communications and Networks Workshops, 2008. SECON Workshops '08, 16-20 June 2008.

CHAPTER 5
REAL USE CASE SCENARIOS BASED ON P1900.4 DRAFT STANDARD

Different use cases are described in the IEEE P1900.4 working group to give more detailed specification of building blocks of IEEE P1900.4 management system. In this chapter, we concentrate on real use case scenarios in order to understand the use of all the functionalities and architectural blocks which have already been described in the previous chapter. We will see via the assistance of well-organized flow charts the exact flow of messages among the entities that have been described in a sententious manner in chapter 4.

5.1 GENERIC PROCEDURES USED TO IMPLEMENT PRIMITIVE USE CASES

Using the proposed system architecture, we present generic procedures for implementing the described scenarios of DS Assignment and DS Access use cases. We propose four following generic procedures: a) Collecting Context Information Procedure, b) DS Assignment Procedure, c) DS Access Procedure, and d) Distributed Radio Resource Usage Optimization Procedure.

5.1.1 Collecting Context Information Procedure

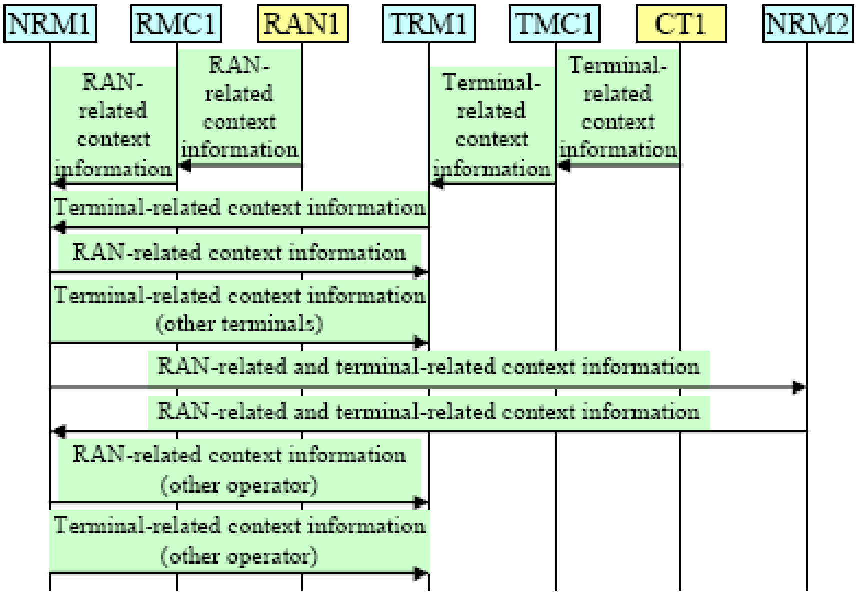


Figure 5.1: Collecting Context Information Procedure [1]

Collecting Context Information Procedure is shown in figure 5.1. RMC obtains RAN-related context information from RAN and provides it to NRM. TMC obtains terminal-related context information from cognitive terminal and provides it to TRM. TRM provides terminal-related context information to NRM. NRM provides RAN-related context information to TRM. Also, NRM may provide terminal-related context information regarding other cognitive terminals to TRM. NRMs belonging to different operators may exchange their RAN-related context information and terminal-related context information. After this exchange, NRM may provide RAN-related context information regarding RANs of other operator to TRM. Also, NRM may provide terminal-related context information regarding cognitive terminals of other operator to TRM. This procedure enables collecting context information required for DS Assignment and Access decisions.

5.1.2 DS Assignment Procedure

DS Assignment Procedure is shown in figure 5.2. OSM generates spectrum assignment policies and sends them to NRM. NRM negotiates regarding DS Assignment with NRMs belonging to other operators (if there is one NRM per several operators, this step is omitted). This last issue is further analyzed in the following section (section 5.2). NRM analyzes spectrum assignment policies and available context information. NRM evaluates current spectrum usage and makes new spectrum assignment decision. Then, NRM informs OSM about this spectrum assignment decision. Eventually, NRM requests RRC to perform corresponding reconfiguration of RAN. RRC controls this reconfiguration of RAN.

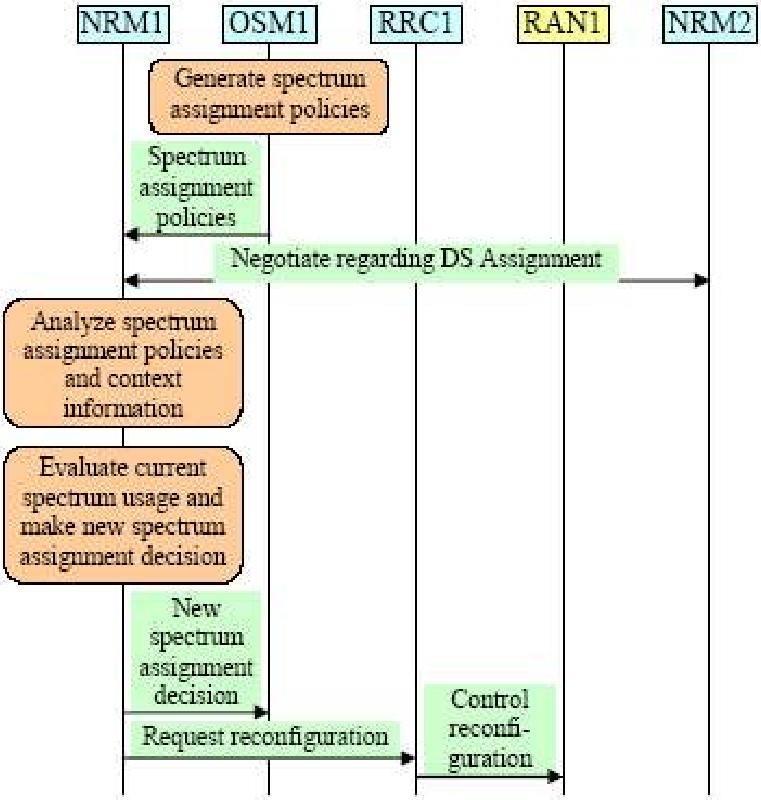


Figure 5.2: DS Assignment Procedure [1]

5.1.3 DS Access Procedure

DS Access Procedure is shown in figure 5.3. NRM analyzes available context information. NRM evaluates current spectrum usage and makes new spectrum access decision. Then, NRM requests RRC to perform corresponding reconfiguration of RAN. RRC controls this reconfiguration of RAN. NRM generates radio resource selection policies and sends them to TRM. TRM analyzes radio resource selection policies and available context information. TRM evaluates current spectrum usage and makes new spectrum access decision. TRM requests TRC to perform corresponding reconfiguration of cognitive terminal. TRC controls this reconfiguration of cognitive terminal.

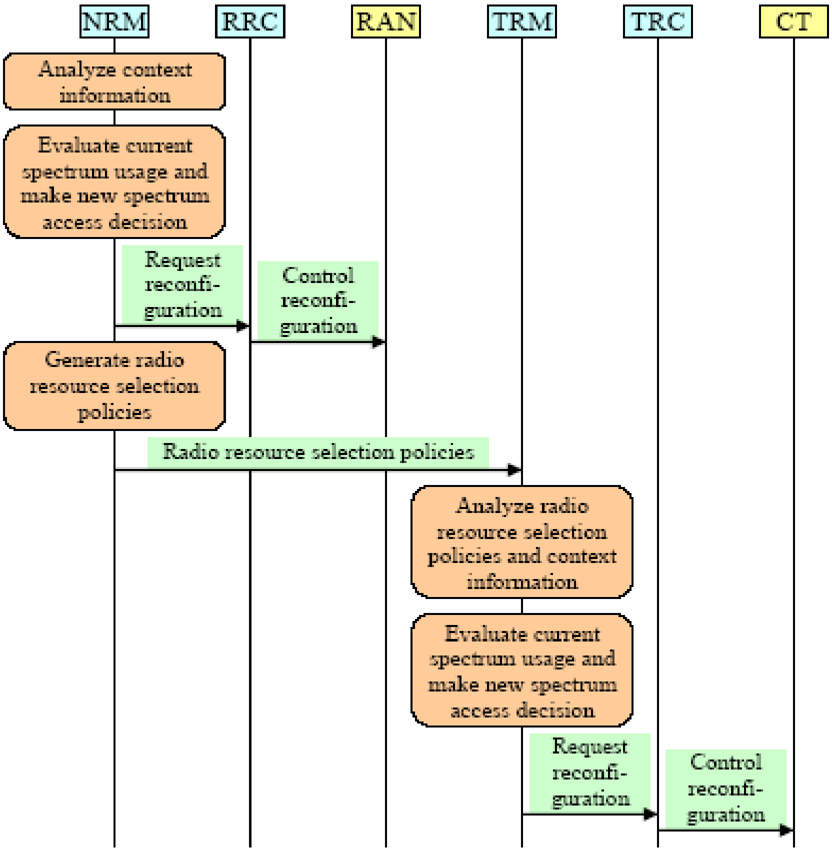


Figure 5.3: DS Access Procedure [1]

5.1.4 Distributed Radio Resource Usage Optimization Procedure

Distributed Radio Resource Usage (DRRUO) procedure describes steps required to make decisions on cognitive terminals reconfiguration in legacy RANs in a distributed manner. It also describes steps required to perform corresponding reconfiguration cognitive terminals. DRRUO procedure does not describe how context information is obtained for this decision making. The whole procedure is demonstrated in figure 5.4.

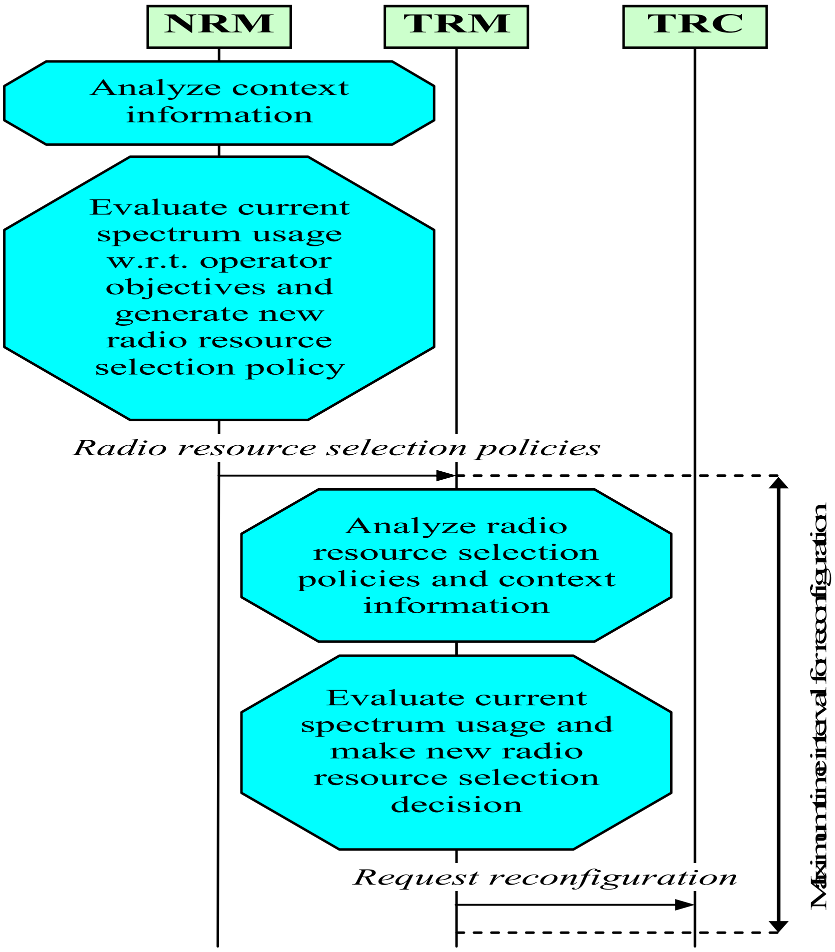


Figure 5.4: Distributed Radio Resource Usage Procedure

More specifically, the exact flow of messages is as follows: 1) NRM analyzes context information, 2) NRM evaluates current spectrum usage and generates radio resource selection policies, 3) NRM sends these radio resource selection policies, 4) after receiving these radio resource selection policies within the specified maximum time interval for reconfiguration (if specified) each TRM performs the following: a) analyzes these radio resource selection policies and available context information, b) evaluates current radio resource usage and makes new radio resource selection decision, c) requests and controls corresponding reconfiguration of cognitive terminal to TRC.

5.2. SINGLE/MULTIPLE OPERATOR SCENARIO REFERRING TO DS ASSIGNMENT

In general, DS Assignment use case refers to the situation when spectrum assignment within heterogeneous wireless networks is dynamically rearranged to improve spectrum usage. In this section, two main scenarios are described providing more details to DS Assignment use case. These scenarios are: a) Single Operator Scenario and b) Multiple Operator Scenario which also includes the optional Multiple Operator Scenario.

5.2.1 Single Operator Scenario

Single operator scenario assumes that several frequency bands are allocated to operator having several RANs. This operator has flexibility for distributing these frequency bands between its RANs. Operator can evaluate efficiency of current spectrum assignment by its OSM. However, it cannot facilitate reconfiguration of its RANs to perform DS Assignment. NRM introduced in the IEEE P1900.4 enables dynamic reconfiguration of RANs inside the operator to improve usage of its frequency bands. Single operator DS Assignment example is shown in figure 5.5. Operator operates RAN1 and RAN2 in two frequency bands. RAN1 frequency band is overused. RAN2 frequency bands is underused. Therefore, spectrum usage is un-balanced. OSM of this operator sends spectrum assignment policy to NRM requesting reconfiguration of its RAN1 and RAN2 networks to allow RAN1 network to use part of RAN2 frequency band. NRM requests and controls corresponding reconfiguration of RAN1 and RAN2 networks. After reconfiguration, part of RAN1 starts to use part of RAN2 frequency band. Usage of frequency bands of operator is now balanced. From the operator's perspective, this allows reconfiguring its networks to balance usage of its frequency bands [2].

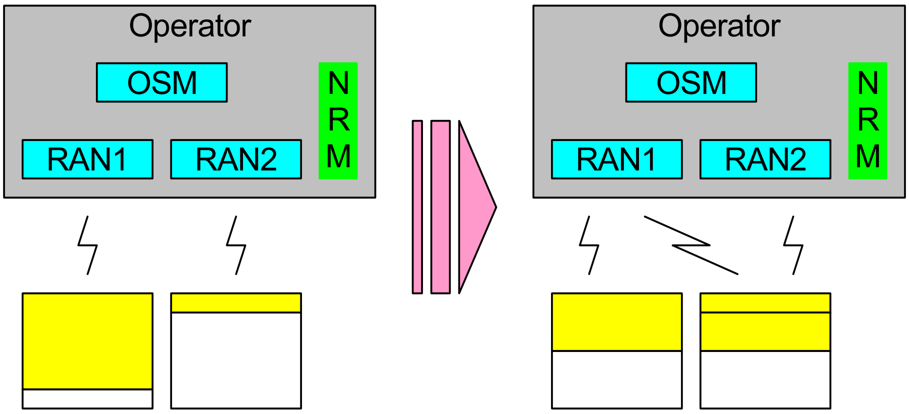


Figure 5.5: DS Assignment: Single Operator Scenario [1]

The exact flow of messages for this particular scenario is as follows: 1) OSM of operator generates spectrum assignment policies and sends them to NRM, 2) NRM of operator receives these spectrum assignment policies, 3) NRM obtains RAN-related context information, 4) TRMs obtain terminal-related context information, 5) NRM and TRMs exchange context information, 6) NRM analyzes spectrum assignment policies and context information, 7) NRM evaluates current spectrum assignment inside operator and makes new spectrum assignment decision, 8) NRM informs OSM about its spectrum assignment decision, and 9) NRM requests and controls corresponding reconfiguration of RANs of operator.

5.2.2 Multiple Operator Scenario

Multiple Operator Scenario assumes that several frequency bands are allocated to several operators and operators have some level of flexibility for renting or sharing these frequency bands. Each operator manages usage of its frequency bands by its OSM. Each operator has its own NRM. These NRMs introduced in the IEEE P1900.4

can negotiate with each other regarding DS Assignment. This enables cross-operator optimization of spectrum usage. Example of this scenario is shown in figure 5.6.

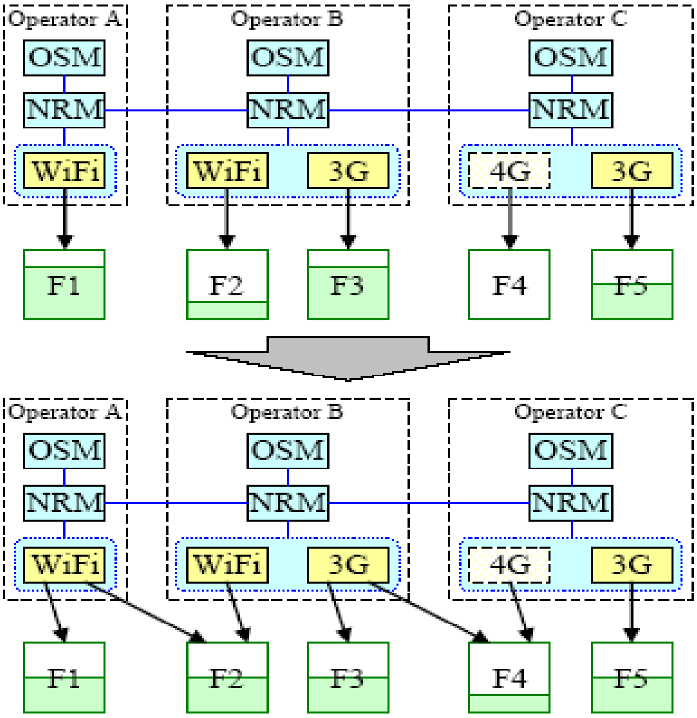


Figure 5.6: DS Assignment: Multiple Operator Scenario [1]

Three operators A, B, and C operate different RANs in different frequency bands. The upper part of figure 5.6 shows spectrum usage before spectrum assignment decision. Operator A operates WiFi network in frequency band F1, which is overused. Operator B operates WiFi network in frequency band F2, which is underused, and 3G network in frequency band F3, which is overused. Operator C operates 3G network in frequency band F5 and has frequency band F4 for deploying 4G network in future. Spectrum usage in frequency bands F1, F2, F3, and F4 is unbalanced. Frequency bands F1 and F3 are overused that degrades quality of service, while frequency bands F2 and F4 are underused, which reduces return of investment.

Improving this unbalanced situation requires cross-operator spectrum assignment decisions, which can be made by negotiation and further decisions of NRMs. OSMs send spectrum assignment policies to their NRMs. Operator A expresses need for additional spectrum for its WiFi network, operator B expresses need for additional spectrum for its 3G network and intention to rent or share its frequency band F2, while operator C express intention to rent or share its frequency band F4.

Based on these spectrum assignment policies NRMs negotiate with each other making new spectrum assignment decisions. They assign frequency band F2 for shared usage by WiFi networks of operators A and B. Also, NRMs assign part of frequency band F4 for usage by 3G network of operator B on the rent basis. After the decisions, NRMs perform corresponding reconfiguration of RANs. The lower part of figure 5.6 shows spectrum usage after the reconfiguration. Now spectrum usage is balanced [1].

Optional Multiple Operator Scenario is a modification of Multiple Operator Scenario. In Optional Multiple Operator Scenario it is assumed that each operator does not have its own NRM. Instead, NRM may be provided by third party. The whole procedure in this scenario is similar and thus we won't analyze it in an extensive manner.

5.3 DS ACCESS SCENARIOS

In general, DS Access use case refers to the situation when RANs and cognitive terminals dynamically access different frequency bands assigned to heterogeneous wireless networks. In this section two scenarios are described providing more details to DS Access use case. These scenarios are: a) Reconfiguration of RANs and Cognitive Terminals Scenario and b) Reconfiguration of Cognitive Terminal with Multi-Homing Capability Scenario.

5.3.1 Reconfiguration of RANs and Cognitive Terminals Scenario

Reconfiguration of RANs and Cognitive Terminals Scenario assumes that one or several frequency bands are available for joint use by several RANs. These RANs and their cognitive terminals dynamically access these frequency bands for improving spectrum usage and quality of service. Decisions on this DS Access are jointly made by NRMs and TRMs in a distributed manner. After the decisions have been made, NRMs facilitates corresponding reconfiguration of their RANs and TRMs facilitate corresponding reconfiguration of their cognitive terminals. An example of this scenario is shown in figure 5.7.

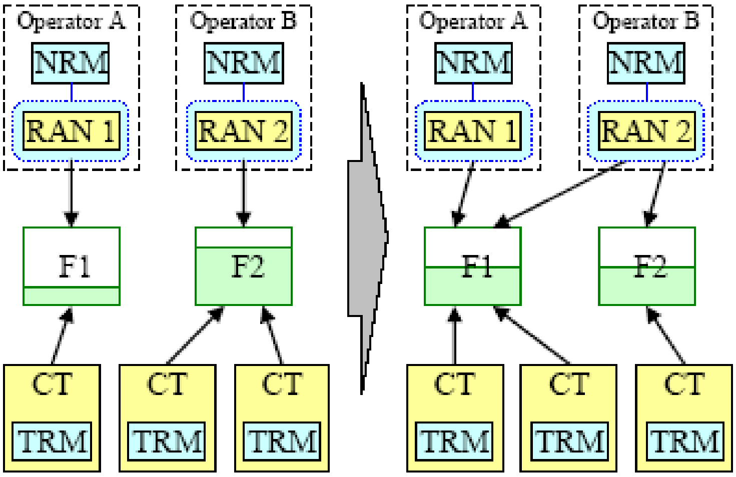


Figure 5.7: DS Access: Frequency band shared by several RANs scenario [1]

Operators A and B operate RAN 1 and RAN 2 in frequency bands F1 and F2. These frequency bands are available for joint use by both RANs. When NRMs and TRMs detect that frequency band F1 is underused, while frequency band F2 is overused, they make decision to shift part of RAN 2 and its cognitive terminals to frequency band F1. Based on this decision, NRM of operator B requests and controls reconfiguration of RAN 2, while TRMs request and control reconfiguration of their cognitive terminals. After the reconfiguration, RAN 2 and their cognitive terminals operate in both frequency bands F1 and F2. As a result, spectrum usage becomes balanced [1].

Conclusively, the exact flow of messages given in an abstract manner for this specific scenario is as follows: 1) NRMs obtain RAN-related context information, 2) TRMs obtain terminal-related context information, 3) NRMs and TRMs exchange context information, 4) NRMs exchange context information, 5) NRMs analyze context information, 6) NRMs make decisions on spectrum access to improve spectrum usage and quality of service, 7) NRMs request and control corresponding reconfiguration of RANs, 8) NRMs generate radio resource selection policies and send them to their TRMs, 9) TRMs analyze received radio resource selection policies and available context information, 10) TRMs make decisions on spectrum access to improve spectrum usage and quality of service, and 11) TRMs request and control corresponding reconfiguration of their Cognitive Terminals [2].

5.3.2 Reconfiguration of CT with Multi-Homing Capability Scenario

Reconfiguration of Cognitive Terminal with Multi- Homing Capability Scenario considers cognitive terminals with multi-homing capability. Such terminals can have multiple simultaneous connections with different RANs. Also, they can dynamically change active connections within the DS Access procedure. Decision on such DS Access is made by TRM of the cognitive terminal and is supported by NRM. This scenario is illustrated in figure 5.8.

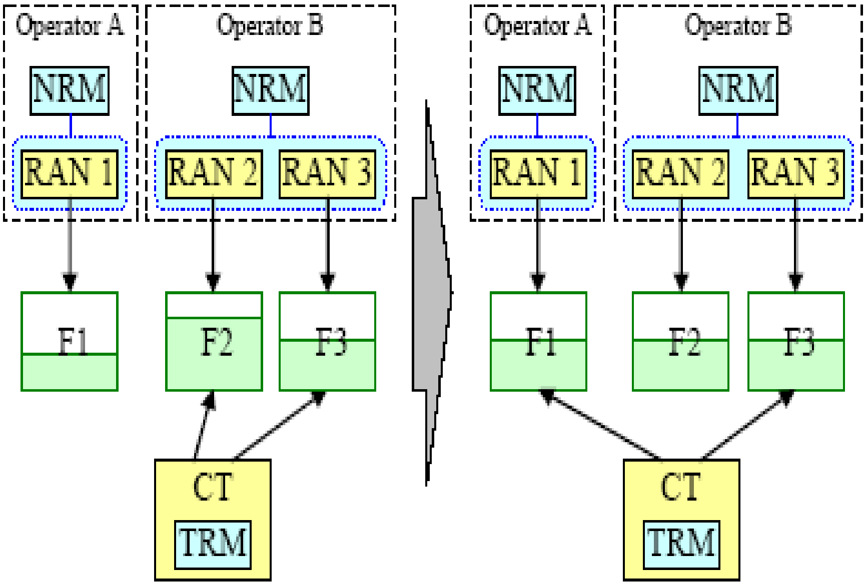


Figure 5.8: Reconfiguration of Cognitive Terminal with Multi- Homing Capability Scenario [1]

Operator A operates RAN 1 in frequency band F1, while operator B operates RAN 2 and RAN 3 in frequency bands F2 and F3 correspondingly. When NRMs and TRM detect that frequency band F1 is underused, while frequency band F2 is overused, they make decision to shift one connection of cognitive terminal from frequency band F2 to frequency band F1. Based on this decision, TRM requests and controls reconfiguration of its cognitive terminal. As a result, spectrum usage becomes balanced.

Conclusively, cognitive terminal with multi-homing capability scenario is enabled by NRM, TRMs, and collaboration between NRM and TRMs. NRM and TRMs obtain and exchange context information used for decision making. TRMs make final decisions on spectrum access for its cognitive terminal, as well as, request and control corresponding reconfiguration of its cognitive terminal. Cognitive terminal with multi-homing capability scenario improves spectrum usage and increases quality of service.

5.4 REFERENCES

[1] Stanislav Filin, Kentaro Ishizu et al, "**Dynamic Spectrum Assignment and Access Scenarios, System Architecture, Functional Architecture and Procedures for IEEE P1900.4 Management System**", 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom, 15-17 May 2008.

[2] IEEE P1900.4/D1.5, "**Draft Standard for Architectural building blocks enabling network-device distributed decision making for optimized radio resource usage in heterogeneous wireless access networks**", April 2008.

CHAPTER 6
802.21 & P1900.4 STANDARDS COEXISTING IN A FUTURE NETWORK ARCHITECTURE

6.1 MOTIVATION OF OUR RESEARCH

In the present thesis, we have already covered a great range of the upcoming topics in the, so called, 4G era. We began with some basic conceptions dealing with the wireless network technologies, which are competing in the next generation era. In the first chapter, the reader can find all the necessary information about the most popular wireless technologies existing nowadays, all being described in a very sententious manner. The research community has started working on all the possible integration aspects of all these heterogeneous wireless network technologies since 2002. A lot of papers have been published since then and a great number of projects are still running worldwide. In our undergraduate thesis, we tried, among others, to make a survey of all the proposed interworking architectures being referred in the international literature until 2007. Our analysis finally focused on UMTS/WLAN interworking paradigm, but it can be easily adjusted in any other pair of networks such those we have presented in chapter 1. Beginning from 2006, a new research trend has emerged and this was the first motivation of our postgraduate thesis. This trend focuses on the supportability of mobility in interworking networks. That is, we want to see how we can achieve seamless vertical handover procedures with the least trade-off.

New draft standards appeared in the last couple of years. IEEE 802.21 appears to be the most promising one facilitating vertical handovers between IEEE 802 and non-IEEE 802 access networks (e.g. cellular networks) in a way that is independent from particular access network features. This standard defines a set of new modules and entities which are going to exchange an even larger set of signalling messages. The final goal is to achieve effective and efficient vertical handovers. The features of the standard are being thoroughly given in a strictly structured way in the second chapter of the present thesis. Chapter 3 is a logical continuance of chapter 2 giving specific examples in order the reader to understand how all the aforementioned messages are delivered in real interworking paradigms.

Despite the fact that there are plenty of open topics in mobility issues of 4G networks, we decided to focus our research interest in a different way. IEEE P1900.4 draft standard was the most important reason for this fact. The reader can identify some aspects of P1900.4 (chapters 4 and 5) that are very familiar taking simultaneously into consideration the features of 802.21. With a more careful investigation, we found out that these two draft standards can complement each other so that we will be able to accomplish optimum vertical handover procedures (802.21) as well as making optimum radio resource usage of the available spectrum (P1900.4).

As a matter of fact, we all know that the wireless resources are finite and can be very limited, especially when we are discussing about a global wireless system which will have the capability of accomplishing goals such as the so called whenever, wherever and whatever services. At this point, we want to clarify that we believe that optimum radio resource usage of the available spectrum is a prerequisite of the whole research endeavour. That is, what would be the necessity of building and implementing the optimum 802.21 architecture with the least delays during the handover procedures, if the system is not capable of providing the necessary radio resources?

In the real systems of the worldwide market operating nowadays, there is certainly a well designed management system which is responsible, among others, to manage the radio resources (RRM modules). As one can understand, this is by no means adequate in the 4G era, as we are dealing with several heterogeneous wireless network technologies, which have to be managed in an even abstract and global way. We believe that IEEE P1900.4 can play this role very successfully. The problem is that the two standards began their routes from different starting points. That is, the members of these two groups did not take into consideration the whole problem, as they want to solve more specific ones. Nevertheless, they have both let open issues, which can be complemented with each other. For example, 802.21 defines the media independent information function (MIIS) just for exchanging information with the MIIS server, letting open issues for other standards, referring that they are outside of the scope of 802.21.

From all the above, it is obvious that there is much space for 802.21 and P1900.4 coexistence in a next generation network architecture. This proposed architecture is presented in the following section.

6.2 THE PROPOSED ARCHITECTURE

In this section, we are trying to present an architecture which integrates the modules and the entities of both draft standards that have already been investigated in the previous chapters. Following the concepts of our research motivations (subsection 6.1), we give an example of the indicant location of the most important entities that have been discussed in the present thesis.

Sententiously, our first objective is to locate the most important entities in an overall next generation heterogeneous network architecture (see figure 6.1). Our second goal is to describe the way the heterogeneous entities will communicate with each other. That is, we will describe the kind of messages that are being exchanged in order to particularize our problem, according to those we have already discussed in the previous chapters. The translation patterns that have to be used so that the entities will have the capability to "understand" each other are out of the scope of this work. Finally, we will explain the exact procedures which are taking place in order to have the capability to accomplish optimum vertical handover procedures as well as making optimum radio resource usage of the available spectrum.

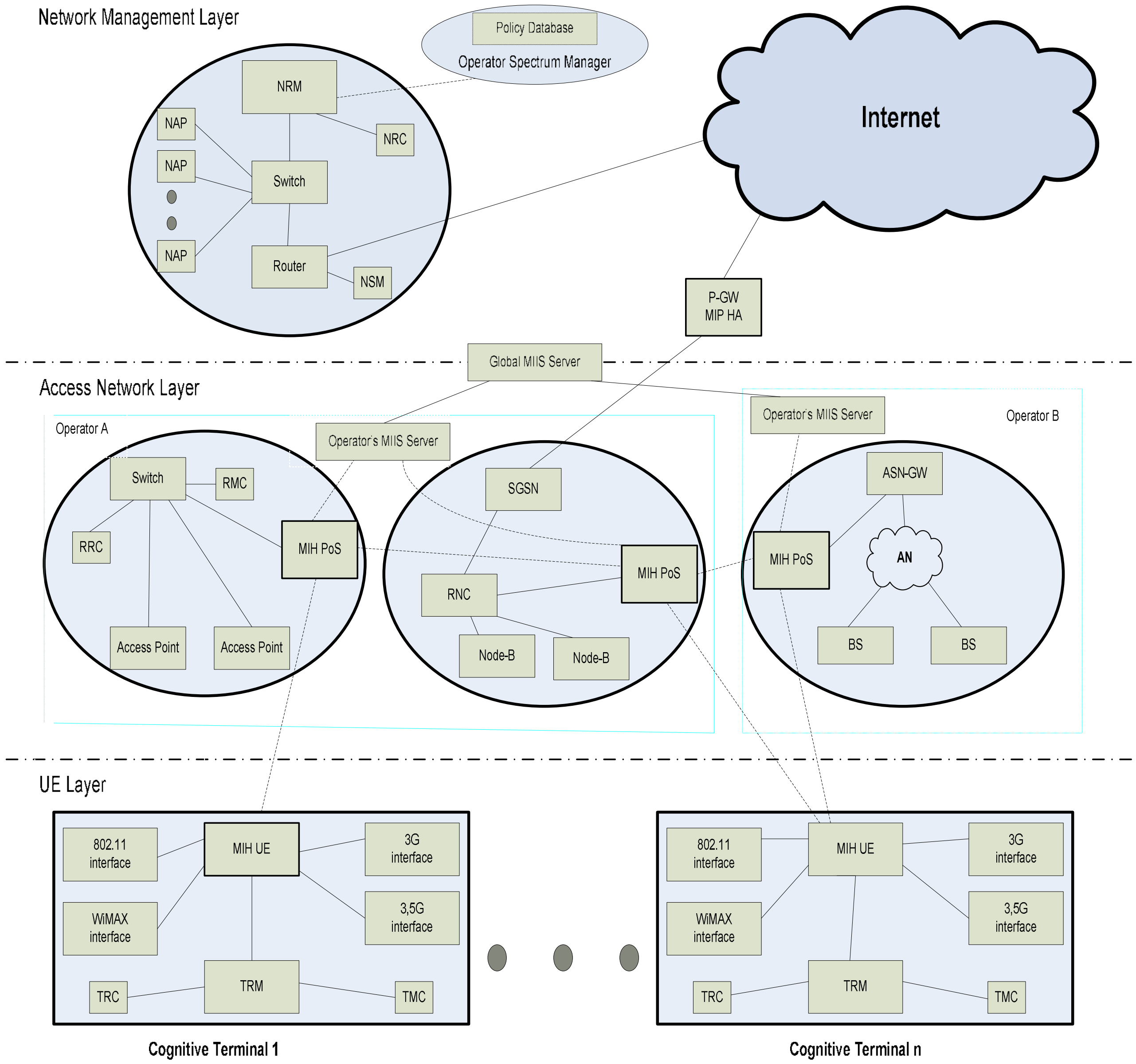


Figure 6.1: The proposed architecture

As one can see in figure 6.1, the whole architecture is divided in three layers, the user equipment, the access network and the network management layer. In the following subsections, we try to present each layer separately, emphasizing simultaneously on the way these three layers communicate with each other.

6.2.1 The User Equipment (UE) layer

At the bottom of figure 6.1, we see the user equipment layer. Here, we find all the cognitive terminals of the future that are supposed to be designed according to the upcoming standards. That is, we suppose that these kind of terminals will have the capability to communicate with the rest network and to understand all the types of messages which are delivered. These terminals can be future mobile phones, PDAs, laptops etc. Conclusively, we are talking about a scenario which cannot be implemented in the real market at the time being but could be a generic solution for the near future.

In figure 6.1, we present only two cognitive terminals for simplicity reasons. It is obvious that whatever is valid for one terminal, the same will go for the others. The main entity one can see inside the terminal is the Terminal Reconfiguration Manager (TRM). Discussing about the P1900.4's functional architecture in chapter 4, we have given an extensive analysis of the sub-modules of this particular entity. In the following figure (figure 6.2) we see the structure of TRM in a very abstract manner.

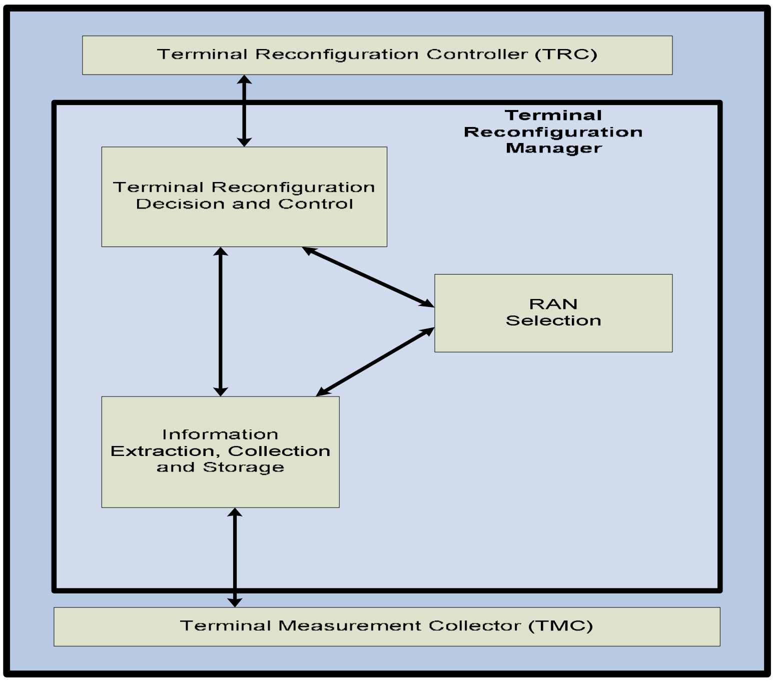


Figure 6.2: The Terminal Reconfiguration Manager

The presence of TRM entity in the cognitive terminal is part of a revolutionary theory which has already been made in the international research community. The new feature is that the user equipment participates in a very active way in various decisions of the overall system. As we have already said, P1900.4 asserts that the final decision about a possible vertical handover will be made by the cognitive terminals. Substantially, what is going to be done is that the terminal will act within a specific framework given by the centralized network system. This framework is supposed to be quickly and dynamically adjusted/reconfigured and that's the main point we are dealing with. What are the messages that have to be exchanged in order the terminal to have the best knowledge about the conditions in the overall network? What should be the right proportion of responsibility in terms of decision matters between the terminal and the network side? These are very challenging questions with which we are dealing with in the present thesis and could be a great challenge for our future research efforts.

The other main entity in this layer is the MIH UE entity, which was thoroughly analysed in chapter 2. The key question here is what is being transferred between MIH UE and TRM? Going back in chapter 3, we can have the necessary feedback for answering this question. More specifically, when a query about the available resources takes place from the serving network to the possible candidate networks, a response is given via a MIH_Candidate_Query.response. This message contains the necessary information for TRM in order the latter to make the appropriate network selection via the RAN selection sub-module presented in figure 6.2. Conclusively, we can say that TRM undertakes the responsibility for making the optimum network selection and MIH UE comprises the basic interface for exchanging information with the access network layer.

Finally, the necessity of the various interfaces presented in figure 6.1, such as the 802.11, the 3G, the WiMAX interface etc is quite obvious. They basically are used for link detections. For example, if a UE enters a WLAN area, beacons from the corresponding Access Point are being "captured" by the 802.11 interface which then transfers the corresponding message to MIH UE entity (Link_Detected and MIH_Link_Detected indications). For more details the reader can refresh his memory going back to chapters 2 and 3.

6.2.2 The Access Network layer

In the middle of figure 6.1 we can see the access network layer. Here, we can see all the available heterogeneous wireless access network technologies within a specific geographical area. It is assumed that there are going to be multiple operators. This assumption does not mean that there is no possibility of single operator scenario. But with the today's market status, we consider the most general version for our 4G network architecture. Thus, in our paradigm we have two operators, operator A and operator B. The first operator is assumed to handle an 802.11 and a 3G network, while the second operator is assumed to be a WiMax operator. At this point, we want to clarify that the selection of these networks is random and the architecture could be much more different. We could also add more operators but it is obvious that our main concern is focused on investigating the main aspects coming from the coexistence of 802.21 and P1900.4 draft standards.

The main entity in the access networks' clouds is the MIH PoS, for which much discussion has been made in chapters 2 and 3. The MIH PoS undertakes the responsibility of being the key connector between the network and the terminal. For example, significant types of messages referring to vertical handover procedures such as MIH_MN_HO_Candidate_Query, MIH_Net_HO_Candidate_Query, MIH_MN_

HO_Commit etc. are being exchanged between MIH PoS and MIH UE entities. The MIH PoSs of different networks and even different operators are also in constant communication state. It is underlined that some of the architecture's connections are omitted due to lack of space.

A constant information exchange is also taking place between the MIH PoS and the operator's MIIS Server. It is assumed that every operator has its own MIIS Server for fulfilling its own needs. There is also a global MIIS Server, which undertakes the responsibility of delivering inter-operator information; that is media independent information services between different operators. The main message which is delivered between the MIH PoSs and the MIIS Servers is the MIH_Get_Information requests and responses, which were thoroughly discussed in chapters 2 and 3.

P1900.4's entities participating in this layer are the RAN Reconfiguration Controller (RRC) and the RAN Measurement Collector (RMC), which reside in switches (802.11) or RNCs (3G) etc. These entities take the feedback from the management layer and they reconfigure the available RANs according to the basic network parameters' instant status. For extensive example scenarios referring to DS Access in P1900.4 framework, the reader can go back to section 5.3 of chapter 5 in order to refresh his memory. After RRC entity has made its decisions, the corresponding information is transmitted to the cognitive terminals via the MIH PoS entity. The structure and the exact functionality of this type of messages haven't been discussed in the present thesis, but this research area is going to become our future work.

Finally, according to figure 6.1, access network layer can communicate with network management layer with two alternative ways. The first one is (see operator A) a tight-coupled way via a switch (or a similar network component) to a corresponding switch or router in the higher layer. The second alternative (loose-coupled) is via Internet using P-GW MIP HA. Thus, all data paths from the involved access networks are combined at the packet data network gateway that incorporates functionalities, such as packet filtering, interception, charging and IP address allocation. Then, P-GW routes traffic to the operator's network or to an external network, such as Internet. We investigate all the aspects of the possible integration architectures (tight, loose, hybrid coupling etc) in our undergraduate work, whose reference lies in the corresponding section of chapter 1. In this thesis, we have not chosen specific integration architecture. Nevertheless, we have included the two most important of them in a single abstract architecture in order to emphasize the alternatives that are given for the information to be exchanged between the two higher layers of our proposed architecture.

6.2.3 The Network Management layer

In this layer, all seem to be clearer. It is composed by entities defined only in P1900.4 draft standard. The main entity which can be easily distinguished is the Network Reconfiguration Manager (NRM). NRM communicates with OSM and according to the information it takes, it finally makes the network reconfiguration decision. As we said in subsection 6.2.1, NRM sets a frame of rules that have to be obeyed by the cognitive terminal in order the latter to make the final decision. In the following figure (figure 6.3), we can see the functional structure of NRM entity.

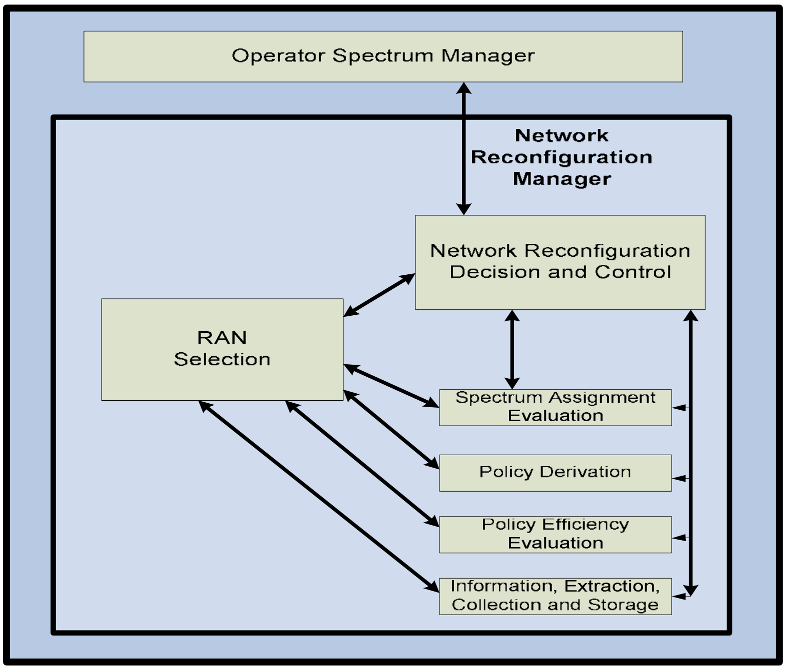


Figure 6.3: The Network Reconfiguration Manager

Another significant entity is the Operator Spectrum Manager (OSM), which generates high level policies having the appropriate feedback about the overall system conditions. An example of a high level policy could be "If the utilization of a 3G network reaches 60%, then generate a trigger for more efficient spectrum assignment". Another one could be "If a UE moves with an average speed up to 2 m/s and it is connected with a 3G network having simultaneously the capability of receiving beacons from a neighbouring AP, then start the handover initialization procedure from 3G to WLAN network". As one may understand, the policies are not that simple but the whole concept approaches the simple policies referred above. NRM gets these policies and tries to make them more specified. For example, it can take into consideration physical or MAC layer information about the condition of the wireless link, various metrics such as PER or BER etc. The result of all the procedure taking place inside NRM is the particularization of the policies generated by OSM and their transformation into more specific commands for the access network and the cognitive terminals.

Finally, NAP is an entity we put on top of the P1900.4 specification. The NAP communicates with a terminal and serves as a node to terminate terminal's multiple links on the network side. The terminal creates multiple IP tunnels between the NAP and the terminal. This aggregation can be used for traffic distribution in order to increase throughput and for duplicated transmission in order to increase reliability. The first case can be very well described via the multi-homing capability scenario presented in subsection 5.3.2 of this work. The second case is also quite interesting as we have the capability to send our packets more than once under certain circumstances. Thus, we can eliminate errors in packets and increase the quality of services received by the end users.

6.2.4 Concluding Remarks

In this section we tried to describe our proposed next generation heterogeneous network architecture in the means of integrating features from both 802.21 and P1900.4 draft standards. Our description was quite abstract as we wanted to give the main guidelines of a system which will have the capability to accomplish optimum vertical handover procedures as well as making optimum radio resource usage of the available spectrum. For this purpose we divided our architecture in three layers described in the previous subsections. The UE layer and the access network layer include the functionalities of the cognitive terminals and the heterogeneous access networks correspondingly. In these two layers, entities from both standards cooperate in such ways that optimum decisions can be made in both network and terminal side. The network management layer includes functionalities that have not been widely applicable in the real market yet.

The main concept that this research work proposes is quite different comparing to those which are taking place in real world market of wireless communications nowadays. Until now, the main concern of the researchers is how to design and implement optimum RRM algorithms in order to optimize the QoS perceived by the end users for each wireless network technology separately. Regarding a next generation global wireless architecture, our main concern should be finding ways for optimum cooperation of the heterogeneous systems. The best way to solve this problem is to have centralized management of the available spectrum.

It is obvious that restrictions in the way the spectrum is allocated to users cannot comprise a good guideline. For example, GSM 900 uses a total of 50 MHz for both uplink and downlink processes. That is, a downlink carrier can use frequencies within 935 and 960 MHz. The new idea is GSM 900 to have the capability of borrowing some other frequencies from a region which belongs to another network in cases of high utilization and congestion. Thus, we will have better overall spectrum utilization which is very crucial in order to accomplish even more bold ideas. For example, according to multi-homing scenario (presented in subsection 5.3.2), the UEs can have the capability of increasing their throughput using the choice of traffic distribution via different networks.

For example let us assume that there is relatively low overall spectrum utilization in the system presented in figure 6.1. The users enjoy relatively good services and everything goes fine. The question is “can we offer even better services to end users?” If we exploit the unused available spectrum, the answer is “yes”. More specifically, in that case a trigger sent by TRMs will be received by OSM in order to change its policies. NRM will undertake the responsibility to give the guidelines of the permitted dealings of spectrum among the heterogeneous subsystems and finally the UE can increase, let’s say, its throughput being simultaneously connected with more than one networks.

As a final conclusion, we can assert that optimum spectrum allocation (in addition with all other entities already existing nowadays) comprises the ideal solution for future’s 4G architecture with the prerequisite that all the necessary messages delivery does not cause excessive problems (such as delays) in the overall system’s architecture.

6.3 FUTURE WORK

We have already made a thorough investigation of 802.21 and P1900.4 draft standards. A lot of discussion has been made about the vertical handover procedures defined in 802.21. Referring to P1900.4, even though we have given all the basic functionalities and specific scenarios, we have not investigated the exact structure of messages which have to be delivered. The last version of the standard created in April 2008 does not include such kind of information. Taking this fact into consideration, we intend to create specific groups of messages, one for each functionality, defining their names, structure, way of transmission etc. Our objective is to build specific flowcharts similar to those presented in chapter 3 for 802.21. Having this first objective fulfilled, we will be able to complete our research in describing the whole system's functionalities.

Our second goal will be to define the heterogeneous messages being exchanged between 802.21 and P1900.4 entities. In this chapter, we have already described the type of information which is transferred among heterogeneous entities in an abstract way. For example, we have referred the type of information being exchanged between MIH UE and TRM. We will continue our research investigating the issue of defining a kind of pattern which has to be adopted in order the heterogeneous messages to be understandable for all the heterogeneous entities.

Having ensured the fact that a smooth flow of information is taking place along the overall system's architecture, we can go on with our final objective; that is, the creation of complete scenarios which can be applicable in the real future market. These scenarios are going to be modelled and plethora of evaluation schemes are going to take place. These evaluations will deal with several quantitative metrics, which will be the criteria of defining the extent of the total improvement which can be achieved with the deployment of the proposed next generation heterogeneous network architecture.