Atheer Dhiaa Al-Rubaye

Vertical Handover Management with Quality of Service Support

Vertical Handover Management with Quality of Service Support

Atheer Dhiaa Al-Rubaye

Universitätsverlag Ilmenau
2017

Impressum

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über http://dnb.d-nb.de abrufbar.

Diese Arbeit hat der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität Ilmenau als Dissertation vorgelegen.

Tag der Einreichung: 6. Dezember 2016
1. Gutachter: Univ.-Prof. Dr. rer. nat. Jochen Seitz (Technische Universität Ilmenau)
2. Gutachter: Univ.-Prof. Dr.-Ing. Kai-Uwe Sattler (Technische Universität Ilmenau)
3. Gutachter: Univ.-Prof. Dr.-Ing. Jochen Schiller (Freie Universität Berlin)
Tag der Verteidigung: 28. Juni 2017

Technische Universität Ilmenau/Universitätsbibliothek
Universitätsverlag Ilmenau
Postfach 10 05 65
98684 Ilmenau
http://www.tu-ilmenau.de/universitaetsverlag

readbox unipress
in der readbox publishing GmbH
Am Hawerkamp 31
48155 Münster
http://unipress.readbox.net

ISBN 978-3-86360-167-6 (Druckausgabe)
URN urn:nbn:de:gbv:ilm1-2017000416

Titelgrafik: photocase.com | AlexFlint

To my parents

Abstract

With a variety of new applications and services offered for mobile users of the Internet, new usage plans and preferences in connectivity to wireless networks might be desired. Connectivity anywhere and anytime through switching between heterogeneous wireless networks became common communication scenarios for many users. To maintain the connectivity for mobile nodes and the continuity of their running sessions, handover decisions, a proper switching scheme between the wireless interfaces of the communication device, and the identification of mobile nodes must be managed. This work presents a vertical handover framework including a mobility management solution as well. It employs multi-criteria decision algorithms that consider a wide range of parameters, mainly to support Quality of Service (QoS) for real-time applications, applies a strategy for stable and soft switching between the multiple interfaces of the mobile device, and presents a light weight signaling scheme for address resolution to quickly recover running sessions. The handover decisions are based on user's configuration, network attributes, and node's context information. A connection is transferred onto a new interface only when it is associated to the newly selected network and ready to take over the traffic. The identity of the mobile node is maintained by leveraging the well-known and widely employed Network Address Translation (NAT) for the purpose of mobility management in a new version that we call Dynamic index NAT (DiNAT). Local and global mobility are supported through hierarchical deployment of DiNAT-enabled anchor points, with no need for pre-knowledge or cooperation of neighbor networks. Many such nodes can be deployed globally for load sharing and route optimization, where a selection mechanism is used to choose a suitable anchor node for each session of a mobile node. The dissertation introduces the proposed approach as a cross-layer system composed of three modules that handle the mentioned tasks, and provides details on the concept of each. The network simulator OMNeT++ is used to model the system and test its feasibility, as compared to a widely adopted solution for mobility management, running real-time applications while moving.

Kurzfassung

Neue Anwendungen und Dienste steigern die Attraktivität der mobilen Nutzung des Internets und fordern die Beibehaltung der Konnektivität auch beim Wechsel zwischen heterogenen drahtlosen Zugangsnetzen, wobei viele Informationen unterschiedlicher Quellen berücksichtigt werden müssen. Auf Basis dieser Informationen müssen Handover-Entscheidungen getroffen werden, die ein Umschalten zwischen den drahtlosen Schnittstellen bewirken und die Identifikation des mobilen Knotens aktualisieren. Die vorliegende Arbeit stellt ein Rahmenwerk für vertikalen Handover vor, das zudem eine Mobilitätsunterstützung beinhaltet. Es verwendet Algorithmen zur multikriteriellen Entscheidung, die eine breite Reihe von Parametern betrachtet, um so die Kommunikationsdienstgüte (Quality of Service, QoS) für Echtzeitanwendungen bereitzustellen. Darüber hinaus wurde eine Strategie für die stabile und weiche Umschaltung zwischen verschiedenen Schnittstellen des mobilen Geräts entwickelt und eine leichtgewichtige Signalisierung für die Adressauflösung zur schnellen Wiederaufnahme der Datenübertragung vorgeschlagen. Die Dissertation beschreibt den schichtenübergreifenden Handover-Ansatz in drei Modulen, deren Konzept und Funktionalität detailliert diskutiert werden. Handover-Entscheidungen werden auf Grundlage von Benutzerpräferenzen, Netzwerkeigenschaften und Kontextinformationen des mobilen Endgeräts getroffen. Eine Verbindung wird nur dann auf eine neue Schnittstelle umgestellt, wenn diese mit dem neu gewählten Netzwerk in Verbindung steht und entsprechend konfiguriert ist. Für die Aktualisierung der Identität des mobilen Knotens wird der bekannte Mechanismus „Network Address Translation“ (NAT) wesentlich erweitert, was als Dynamic index NAT (DiNAT) bezeichnet wird. Sowohl lokale als auch globale Mobilität werden durch eine hierarchische Bereitstellung von DiNAT-fähigen Knoten unterstützt, ohne dass hierzu ein Vorwissen oder die Kooperation der Nachbar-Netzwerke notwendig ist. Viele solcher Knoten können zur Lastverteilung installiert werden, da die Dissertation einen Auswahlmechanismus erarbeitete, um die Auswahl eines passenden Ankerknotens für jede Kommunikationssitzung auf einem mobilen Knoten zu ermöglichen. Die Simulationsumgebung OMNeT++ wurde verwendet, um die Durchführbarkeit des Ansatzes zu überprüfen und ihn mit einer weit verbreiteten Lösung für das Mobilitätsmanagement zu vergleichen.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor Professor Jochen Seitz for his guidance, patience and encouragement. His support was very important along the time of my research work.I would like to thank Professor Kai-Uwe Sattler and Professor Jochen Schiller for reviewing my thesis, and the committee members as well for their insightful comments. I feel also very grateful to Mrs. Gudrun Breternitz from the international consulting office at TU Ilmenau for her support. Thanks to the Deutscher Akadimischer Austauschdienst (DAAD) for providing me the scholarship. I am especially thankful to my parents who supported me spiritually and always prayed for my success. Thanks to my father for believing in me. Thanks to my mother for her love. She would have been so happy to see this accomplishment. Many thanks to my lovely family; my wife Zeena, who shared me the highs and the lows along this journey, and my two sons Hayder and Hussain, whom existence and smile always inspire me and give me the determination to achieve my goals.

Contents

1 Introduction

In today's emerging communication networks, mobility and switching of connectivity in between, known as handover, while having applications running on a smart device became a part of everyday's usage scenario for so many users. The different wireless networks surrounding us represent heterogeneous systems that differ in their attributes and features. They usually represent different subnets and belong to different administrations. Therefore, handover between such networks, known as vertical handover, is a more challenging scenario than limited mobility within the same network. A primary characteristic of next generation networks (i.e. 5G and beyond) will be the integration of various radio access technologies to provide unbroken connectivity and services anywhere and anytime with optimized and enhanced data rates as well. Heterogeneous networks can be integrated to provide ubiquitous environments and deliver better services when suitable protocols are developed for the current modern and powerful multiple-interfaces smart phones and devices. Therefore, models for handover strategies, decision algorithms, and mobility management should be further developed to adapt to the trending usage.

This dissertation investigates the phases of a vertical handover and describes the challenges of the process. It proposes a solution that provides a sophisticated handover, taking into account the parameters affecting the quality of the provided services, manages executing the process with no need for user interaction, and applies a minimal impact on running applications through quick recovery of traffic. This chapter describes the motivation of this work, addresses general challenges to achieve the desired goals, and summarizes the contribution of the dissertation. The outline of the dissertation is present in the last section.

1.1 Motivation

The notion and nature of user's access to wireless networks and the provided data services and applications for mobile nodes are witnessing a massive development. For many emerging applications and hence, the end users, having connectivity anywhere and anytime, and being always best connected became very common usage plans.

Very interesting statistics and forcasts were recently presented by Cisco in their visual networking index in [Sys16]. It shows the tremendous growth in the size of the global mobile data traffic, which grew 74% to reach 3.7 Exabytes (1 Exabyte = 10^9 Gigabyte) per month in 2015 and is expected to reach 30 Exabytes per month within 2020. The smart phones (including tablets) already represented 97% of the total current global handset traffic. Mobile video traffic accounts now for more than half, and will grow to three-fourths of the total mobile data traffic. Data offloaded from the cellular networks onto the fixed network through Wireless Fidelity (WiFi) networks (realized by handover) represents 51% of total mobile data traffic in 2015 and will increase to 55% in 2020 otherwise, global mobile data traffic would grow to 62% instead of 57%. Impressive amount of traffic is also generated due to wearable devices and device-to-device communications, which may need to handover as well.

Four main reasons for further research in the field of mobility and handover are addressed here, these generally are; the networks, the available devices, the emerging applications and the user demands. In more details:

- The networks nowadays are evolving in terms of connectivity, from lower to higher generations (3G, 3.5G, 4G and 5G). In the upcoming generation, faster Internet connections, higher capacity of users and integration of heterogeneous radio access technologies are to be realized. This leads to a wide adoption of multimedia applications and therefore, the consumption of high data volume per user.
- Smart phones became basically cheaper and available as an average wireless communication device. It is equipped with multiple wireless interfaces for different access technologies. Enhancements on battery life also continue to improve the

utilization of the device's features. Nevertheless, machine-to-machine and Internet of things are new scenarios of communication that can add to mobile traffic and demand a sustained connectivity.

- Many new applications are introducing a new notion of networks utilization. Beyond traditional E-mail and web-based applications, various social, multimedia and automation applications are personalized for mobile nodes, too. Voice and video call applications are examples for the replacement of traditional expensive calling methods with cheap ones.
- With suited environments, devices and applications, users might want to a select among a collection of usage policies suitable to the applications they run at certain times or their preferences of cost and quality.

Anyhow, handover to maintain connectivity and services is becoming a typical communication scenario and will be a very frequent process in daily life. However, the variety of motives for more investigations and revision in this field applies challenging tasks as well.

1.2 Problem Statement

Employing heterogeneous networks requires more than just installing multiple interfaces on wireless devices. The associated challenges can be summarized as follows:

1. Due to the many handover decision preferences from different perspectives, a rich amount of information need to be collected and probably exchanged as well. These might be related to networks parameters, mobile node context information, user preferences and running applications.
2. Heterogeneous networks have different characteristics, so they are not directly comparable when thinking of sophisticated decisions. The decision algorithms need to consider many network parameters simultaneously, with preferences and criteria weights as well.

3. A handover strategy or policy should be designed to manage a stable, seamless and soft handover. Unstable switching back and forth between networks introduces undesired delay and unnecessary signaling with each handover. To have it seamless means to automate the process upon changes in the decision parameters. However, performing the process in a soft way that assures the activity of the second interface before disconnecting the previous one should guarantee the availability of the new connection before actually abandoning the connected one. Such a switching paradigm should conceptually be beneficial in terms of the experienced quality of service.
4. When the mobile node obtains a new address in the new network upon a handover, problems regarding unification of the node identity appears and risks the continuity of running communication sessions. The traffic should be quickly resumed in the new network.
5. Heterogeneous networks are owned usually by different operators and represent various administration domains. Handover in this case applies issues regarding security, cooperation, billing, and compatibility.

1.3 Goals of the Work

As the title of this dissertation states and conforming to the described motivations and challenges in the previous sections, the main goal of this work is to develop a handover model towards providing a ubiquitous access environment over today's coexisting heterogeneous networks, considering users' and applications' requirements and supporting quality of service parameters. In more details, this dissertation aims to:

1. Investigate the phases of the handover process in general and address the requirements of a seamless and soft handover that supports sophisticated decisions and quality of service.
2. Design an overlay module to apply the handover strategy and control the avail-

able wireless interfaces with no regard to the radio access technology underneath.

3. Implement a suitable mathematical decision algorithm that can be inquired by the former module to select between alternatives according to multiple and differently weighted criteria.
4. Implement an address resolution mechanism to handle the mobile node identity issue. Handover delay should be minimized through a light weight signaling scheme.
5. Consider scalability in the proposed model and minimum modifications to already installed networks and running protocols.
6. Model the proposed schemes using a network simulation tool and evaluate them with comparison to other recognized solutions.

1.4 Contributions

The contributions that have been achieved along this dissertation are briefly:

1. A review into handover process and its individual phases, its associated problems and the related known approaches in the field.
2. The development of a new cross-layer handover solution that manages vertical handover including all of its phases. It supports multi-criteria decisions, user's input, soft switching, and manages addressing through enhancements to available basic networking functionalities.
3. The achievement of a quick solution for today's networks. Flexibility in setup, scalability, simplicity and low overhead to network infrastructure and protocols are its main features.

1.5 Outline

This dissertation is organized as follows:

Chapter 2 provides a background on the principle of handover and its classifications. It defines its phases, describes some main types of decision algorithms and addresses requirements for a seamless soft handover. A discussion at the end of this chapter focuses on the main ideas.

Chapter 3 gives an insight into addressing in communication networks and describes some of the recognized mobility management solutions that handle the identity issue of mobile nodes in handover.

In chapter 4, the proposed solution is presented with a detailed description for its modules and their functionalities. An overview on a the handover process adopting the suggested mechanism is also provided including signaling, traffic flow and the process in each of the involved nodes.

The simulation environment, network topologies, and setup scenarios are presented in chapter 5. Details on the simulated scenarios and the obtained results are given and discussed. A brief summary follows at the end of the chapter.

Finally, chapter 6 concludes the work and addresses some open issues for future work.

2 Principles of Handover

As wireless access to the Internet became an integral part of our daily life, continuity of running services while moving introduces new aspects and challenges into connectivity and access to wireless networks. Long-Term Evolution (LTE), Wireless Local Area Network (WLAN), and Universal Mobile Telecommunications System (UMTS) are examples of heterogeneous networks. Figure 2.1 shows a topology of coexistent networks that overlap in many areas. A user might wish to have a persistent connectivity for a running application, to the Internet for example, on the way from home, through the city and to the office.This chapter gives an overview on the handover process, its types and classifications, the decision making and the requirements for a seamless and soft switching between heterogeneous systems like in the shown scenario.

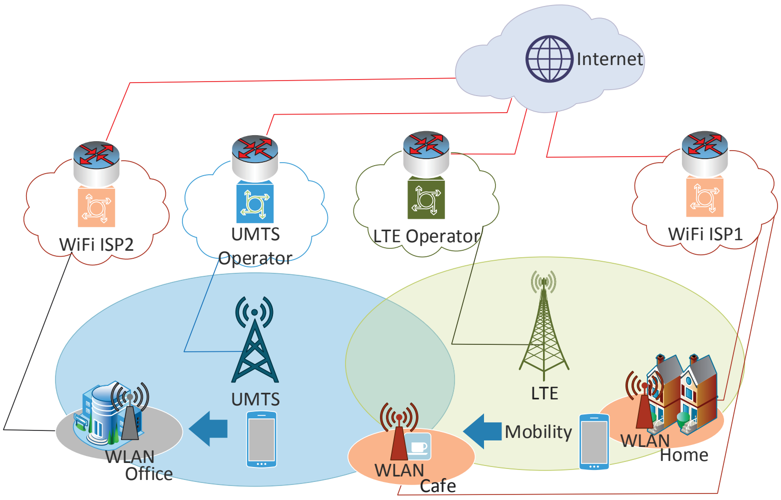

Figure 2.1: A typical scenario of vertical handover.

2.1 Overview and Classifications

The process of switching connectivity from an Access Point (AP) to another is generally known as Handover, which is mainly categorized into horizontal and vertical by the research community. Three main scenarios of wireless access might be generally identified, which can be related to the evolving forms of wireless communications as well, these are; a basic wireless access to avoid wire-line connectivity, in which MNs move only slowly within a range of a single AP, nomadic access, where MNs might switch into another AP when in idle mode (no sessions are running), and mobile access, where MNs might move with no regard to any condition. The last is a very common scenario today, but challenging at the same time since session continuity while moving became a strongly required feature.

When a MN moves between two access points/cells of the same radio technology, it is defined then as Horizontal Handover (HHO). A handover between access point of different radio access technology is referred to as Vertical Handover (VHO). Sometimes, another type is farther defined as diagonal handover, which is the case when access points belong to the same underlying technology, for instance, the IEEE family [ABG14].

Besides the type of the radio access technology, a handover can be further classified based on other criteria, such as the number of the involved access points, the administration domains, the initiator of the handover, terminal state, and service continuity [Wol00, KKP08].

Based on the number of involved APs, the HO can be classified into hard and soft. The former is also referred to as Break-before-make, where a MN can be in connection to only one AP a time. The later is referred to as make-before-break as well, where two APs might be in connection to a MN during the handover. This type, however, requires overlap in coverage areas. A third type can be defined when a MN switches between radio links that belong to the same AP and this is referred to as softer HO. [DS14, NHH06].

A handover is considered as local or Intra-domain when a MN moves within the same administration domain, while referred to as global or Inter-domain when moving between different domains. Networks of different access technologies usually represent different administrations, since they might be owned by different operators, and therefore, a VHO represents in most of the cases an Inter-domain handover as well. Similarly, handover can be classified into Intra-/Inter-cell HO, Intra-/Inter-BS controller and Intra-/Inter-switching center [Sch03], which are common terms in horizontal handover within cellular systems.

Considering the initiation and decision making entity of a handover, it can be categorized as terminal-controlled and network-controlled. However, when measurements and information are collected by the mobile, but used by the network to decide and control a HO, it is referred to as network-controlled mobile-assisted HO, and vise versa, when the network collects information that can be used by the mobile node to make a HO decision, it is called then mobile-controlled network-assisted HO.

Different networks are identified by a variety of subnet addresses and therefore, a MN might obtain a new IP address in consequence to handover. A handover is said to be a L2 handover when switching between APs that belong to the same subnet address, where no change in the MN's IP address follows. A L3 handover includes changing the IP address, where the APs belong to different subnet addresses.

Taking service continuity into account, it can be named as seamless and non-seamless handover. In the former, a handover is unnoticeable to the user, while a user can notice the last.

The reasons to initiate a handover can be considered to define two further classes; imperative or forced handover, which is triggered by events related to signal availability, and alternative or user handover, where polices and user preferences play the major role in the HO decision [KKP08].

Heterogeneous networks (in the sense of radio access technology) can be coupled in a tight way, in which all the access points share the same subnet, so handover would be easier in management in such a topology since a MN owns the same IP address before and after switching. This setup assumes a centralized core network but different radio technologies at the access level. On the other hand, loosely coupled networks have different subnets, which are reachable for each other.

According to the earlier definitions, this work addresses mobile-controlled network-assisted seamless soft vertical handover process between loosely coupled IP-based mobile wireless networks. The reasons behind this assumption are:

- Introducing no extra processing load on the network side through the development of mobile-controlled solutions, where frequent handovers are expected to increase due to the pursue of almost permanent connection status. However, network-assisted approaches can exploit the provided network-related parameters, like the available capabilities and services of a network, as useful information to the MN's HO system.
- Reducing packet loss and making the switching unaware for the running real-time applications, and hence, the user, as a basic measurement of quality.
- A simple scenario would be to handle layer 2-limited handover between tightly coupled radio access networks, in which they are connected in a single core. However, today's realistic topologies apply scenarios where each network operator is connected to the Internet through his own gateways and has his own independent administration and ownership of network equipment. Therefore, loosely coupling of the heterogeneous networks as described enforces layer 3 to be involved in the handover process.

2.2 Handover Process

Handover is an essential process in modern mobile networks and usage scenarios. According to [KKP08, SV09, ABG14], any handover includes three main phases:

- **Measurements and network discovery:** During this phase, measurements of HO-related parameters are carried out by the MN (and/or the AP). A MN determines which networks can be used and their related capabilities that might be advertised.

- **Decision Making:** A handover process might be initiated when the collected parameters meet some predefined conditions, threshold levels for example. The objective of this phase is to select a better alternative connection and the exact time to carry on a handover. A mathematical decision algorithm delivers the result of comparing available alternatives. Regardless of the mathematical analysis, the handover decision might be controlled by a HO management policy, which determines if handover should be carried out and when exactly.

- **Execution:** In this phase, a MN transfers its connection to the new selected AP/network. It might involve link association, node reconfiguration, database lookups, signaling and authentication procedures as well. Redefinition or managing address resolution for MNs to resume traffic is triggered by the completion of this phase, and is often referred to as mobility management in the field of wireless communication.

Along the literature of handover, the research work considers these phases with different perspectives and focus. The main efforts on handover usually focuses on the decision phase rather than on network discovery, and the mechanism used to execute a handover, if decided. These two phases are explored deeper in the horizontal handover research from the prospect of channels allocation and management of radio resources.

However, simple handover approaches normally consider a single parameter or two, which in many cases are available and easy to measure in the device, to take the handover decision. A framework called Media Independent Handover (MIH) is often employed for more sophisticated approaches to assume the availability of the required parameters and information. On the other side, work on mobility management focuses on handling issues that arise in a MN from carrying out a handover.

Accordingly, handover decision algorithms are discussed more in details than the other phases in this chapter, while mobility management is deeply investigated in the next chapter.

Media Independent Handover

IEEE 802.21 MIH is a framework to collect and exchange handover-related information. It represents an under-development approach for the deployment of handover services over heterogeneous networks [MIH09, OBS+08]. MIH encourages cooperative use of information available at the mobile node and within the network infrastructure. The MIH protocol stack should be implemented in all the devices involved in the handover if to be realized. Essentially, it consists of a framework, a set of handover-enabling functions (MIH Functions - MIHF), and an MIH Service Access Point (MIH SAP). The MIHF provides the Media-Independent Event Service (MIES), the Media-Independent Information Service (MIIS), and the Media-Independent Command Service (MICS). In general, the MIHF provides mobility management system services to MIH users (MNs and network nodes) through the MIH SAP and interfaces (i.e., receives events and send commands) with lower layers through media specific SAPs, as figure 2.2 illustrates [MIH09]. As a logical entity, MIHF is also able to exchange messages with other MIHF peers about access network status in the area surrounding the mobile node.

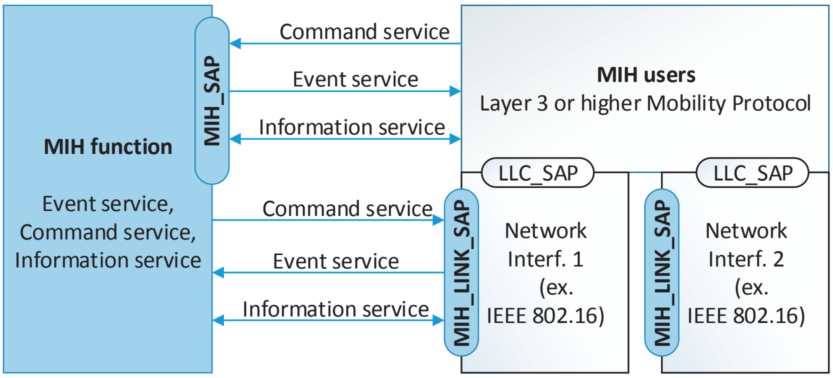

Figure 2.2: The MIH services

However, MIH is only a signaling scheme to support handovers. It does not implement a mobility management solution neither a selection mechanism. It is a general framework to exchange information and commands useful to be deployed in VHO solutions. Furthermore, its paradigm of communications with nodes/servers adds supplementary signaling and high resource consumption, despite that there were efforts done to enhance the information server architecture [KPKP11, NSS+11].

The authors in [JCV+13] extended MIH with entities for quality of experience estimation, mapping and adaptation insurance. MIH entities installed on the MN and the base stations/access point as well. When a MN detects a target network, it inquiries for the network resources information through MIH. A handover is decided if the estimated quality of the running video traffic in the targeted network is better than in the connected one when using one of the mapped quality classes. The adaptation mechanism can adapt the application according to the network conditions by dropping less important packets in cases of congestion.

[LL11] presents a handover approach using MIH with a location-based architecture. A MN sends a candidate request, which contains the MN identity and location, to its MIIS server when a link goes down. The MIIS finds a suitable AP for the MN according to its position. MIIS servers are deployed hierarchically in this approach. A lower MIIS may inquire a higher one when it has no suitable candidate for a certain MN request. A MN sends a binding update to a location server in its home network in case of a HO. When a node wants to communicate with a MN, it contacts the location server first to receive the current location of the MN. Such an approach applies much efforts and signaling to identify the target AP, which adds in turn delay to HO latency and traffic recovery.

In [LSK+09], a vertical handover employing the MIH is proposed. A main goal is to provide a balanced load among the networks. The solution is limited to mobility within a single administration domain.

2.3 Handover Strategies in our Smartphones

Many different brands of smartphones in terms of hardware and operating systems manufacturing are currently available in the markets. Various vendors might use different mobility management strategies. However, the connectivity of a smartphone is usually managed by an entity widely known as *Connection Manager*, which decides which network to connect to and how switching should take place.

An Interesting experimental analysis on smartphones is carried out in a recent work by Sanchez in [SdlOB16] to characterize the behavior of the connection manager in three dominant operating systems in the market; Android, iOS and Windows Phone 8. The analysis included a comparison on initial attachments and handovers with real-time applications running.

A common strategy in connection management is to have the cellular connection as a default, while the WiFi as a preferred regardless of the user preferences or connection quality. The only choice left to be under user control is to enable/disable interfaces. Regarding handover, all the examined systems switch to the cellular connection as soon as a WiFi connection fails, even when there are other WiFi access points available. A WiFi connection is recognized as failed only after the loss of several beacons and the fail response to probe requests sent by the system to the access point due to no reachability.

Unlike iOS, Android and Windows Phone 8 found to renew their IP address configuration upon a handover, even when the two networks belong to the same IP subnet. This adds delay to the handover process. When actually switching the connection, iOS does not allow two active interfaces concurrently, while Android keeps the cellular connection active till the WiFi is configured. Windows Phone 8 allows further for simultaneous connectivity, where applications start before switching would continue using the previous interface, while applications newly started use the newly connected one. To enable simultaneous connections, routing tables are implemented per interface rather than system wide.

One of the surprising findings is, however, none of the examined systems uses quality measurements in handover decisions, but rather select the network which was lastly used. Another issue is that mobility solutions are deployed within the same operator network, but there is no own solution implemented in these systems for inter-technology and IP mobility.

2.4 Handover Decision Algorithms

The handover decision is a critical and essential part in every handover approach and has a great influence on its overall performance.

In scenarios of handover between heterogeneous systems, the different network characteristics might not be so directly comparable, therefore, the decision becomes a more complex process. The decision criteria a VHO might consider can be classified as static, like user's preference and service cost, and dynamic, like measured parameters related to; the MN, the application and the network.

In the literature of handover, there exist many classifications of handover decision schemes based on the considered criteria and the methodology used to process them in order to evaluate the alternative networks [KKP08, RAD13, ABG14]. A brief overview on some of the most well-known VHO decision algorithms found in the literature is as follows:

2.4.1 Received Signal Strength based Decisions

The RSS of the current connection is compared to that of another available one in such schemes. The simplicity of the operation and the availability of the parameter made this scheme to be considered by so many handover solutions in earlier work on handover [ABG14, YHI12, YeN10]. A threshold value of RSS is also used such that a network is selected if it satisfies the constraint:

$RSS_{new} > RSS_{old}$ and $RSS_{old} < RSS_{threshold}$

Signal-to-Noise and Interference Ratio (SNIR) is sometimes used with or instead of RSS as a more stable channel indicator [SR11]. SNIR-based algorithms can provide in overall a higher throughput than RSS-based since the available throughput is SNIR-dependent. Due to high variations in SNIR values, such algorithms may produce unstable behavior, where the MN may handover back and forth between the networks [YeN10, Pol96].

2.4.2 Cost Functions based Decisions

A cost for each network out of the alternatives might be calculated according to the running applications preferences. It is calculated as a sum of weighted functions of certain cost or QoS-related parameters [KKP08, ABG14]. The general form of the benefit function f_n in network n is:

$$f_n = \sum_s \sum_i w_{s,i} * c^{n_{j,i}} \tag{2.1}$$

$c^{n_{j,i}}$: is the normalized ith parameter for the service s in network n.
$w_{s,i}$: the weight of the ith parameter to gain the service s.

Including bandwidth, battery power and delay, the network with the minimum cost was selected in [ZM06]. Similarly, in [WKG99], the network calculated to have the lowest cost in terms of bandwidth, power consumption and monetary cost for the maximum services was chosen. Such functions, however, do not consider user satisfaction and node's context information.

2.4.3 User Surplus based Decisions

This is basically a kind of user-defined policy scheme. It evaluates the decision from the user's point of view as the most convenient to his specific needs. For instance, a user might never abandon a cellular access without connection blackouts to ensure quality, even when it costs, or to search for any available WLAN to save cost, even when sacrificing QoS. The authors in [CM04] proposed similar policies, where simulation results showed that the performance of some applications running on the user terminal (FTP, HTTP, and Telnet) was improved whereas others became worse.

2.4.4 Multiple Attributes based Decisions

A VHO represents a typical Multiple Attributes Decision Making (MADM) problem, where a single alternative has to be selected out of many based on different characteristics. Terms such as multiple objectives, multiple attributes, and multiple criteria are used often interchangeably in the literature of decision making. We refer to this scheme as Multiple Criteria Decision Making (MCDM) in this dissertation.

Some of the most popular MCDM algorithms for multi-criteria handover are Simple Additive Weighting (SAW) [Fis67], Weighted Product Model (WPM) [TM89], Analytical Hierarchical Process (AHP) [Saa08], Grey Relational Analysis (GRA) [HI10], VIseKriterijumska Optimizacija Kompromisno Resenje (VIKOR) [OT04], [OT07], Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) [HY81], and Fuzzy MCDM.

In SAW, the score of a network is calculated as the weighted sum of all the attribute values. This model might be referred to in the literature as the Weighted Sum Model (WSM), too. The score of network i is the sum of the normalized metric values $p_{i,j}$ multiplied by the corresponding assigned weight w_j of that metric j. The candidate network with the maximum score is selected, as might be expressed by the following equation:

$$A^*_{SAW} = max_i \sum_j w_j p_{ij} \tag{2.2}$$

where $\sum_J w_j = 1$.

WPM is a similar model. Instead of the sum, the product is used with the weight being in the power, as expressed in next equation. The winner is again the alternative with the highest score.

$$A^*_{WPM} = max_i \prod_j p_{ij}^{w_j} \tag{2.3}$$

In multi-criteria decision problems, besides selecting an appropriate alternative, specifying weights for each criterion might be uncertain task. Both of the former algorithms have no support for calculating weights.

AHP provides a logical framework to solve such decision problems and yields benefits of each alternative analytically. The problems in AHP are decomposed into a hierarchy of the objective, criteria and alternatives. Both qualitative and quantitative criteria can be compared to derive weights and priorities. The relative importance of each criterion is expressed in a pair-wise comparison. Each criterion can be evaluated in terms each of the others criteria in a scale from 1 (equal importance) to 9 (extremely more important). For instance, bit rate might be expressed to be 5 times more important than service fees in a certain policy, and can be therefore, expressed as $bitrate = 5 \times cost$.

A pair-wise matrix is composed accordingly, which can be analyzed to express the normalized weight of each criterion w_j through the evaluation of its eigenvector, where for j criteria, $\sum_j w_j = 1$.

The alternatives can be evaluated over each other in terms of each criterion in the same way, however, quantitatively when it is related to a measurable parameter and qualitatively when related to subjective parameters (user preference for example). A final step includes a multiplication of the weights rank vector by the matrix composed of the alternatives rank vectors (each related to one of the criteria). The result is a vector with values representing the score of each alternative.

For example, the AHP structure would look like as in figure 2.3 when evaluating $N = 3$ of networks using $K = 4$ of criteria, where w_j is the calculated absolute weight of the jth criterion, $nw_{j,i}$ is the score of network i in terms of criterion j, $\sum_j w_j = 1$ and $\sum_i \sum_j nw_{j,i} = 1$.

In GRA, the best alternative is determined by measuring the minimum distance to an ideal solution. This is determined using the Grey relational grade, which is included in the following formula:

$$A^*_{GRA} = min \frac{1}{n} \sum_{j=1}^{n} w_j \zeta_{ij} \tag{2.4}$$

ζ_{ij} is the Grey relational coefficient of the ith alternative according to the jth criterion. It is a representation of the distance to the ideal point and is calculated using further detailed formulas.

The VIKOR algorithm is based on the idea of ideal and compromise solution. It selects the minimum distance to the ideal solution, but in more complex calculations.

In TOPSIS, the basic concept is that the best network should have the shortest distance to the ideal solution and the farthest distance to an anti-ideal solution. The selection formula can be expressed as:

$$A^*_T PS = max_i \frac{d_i^-}{d_i^+ + d_i^-} \tag{2.5}$$

where d_i^+ and d_i^- are distance measurements of the ith alternative from the ideal and the worst (called anti-ideal, too) solutions, respectively.

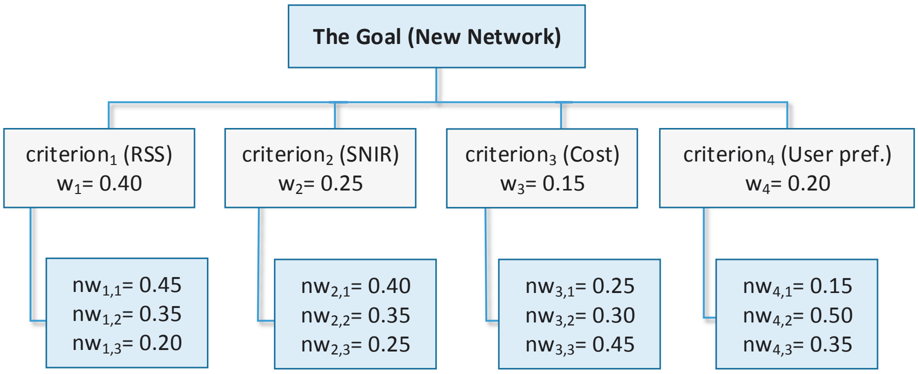

Figure 2.3: The structure of AHP for N of networks in terms of K criteria. The networks are ranked in terms of each criterion j, then a final rank.

However, the distance-based schemes are considered generally as complex algorithms from calculations view [SDD13]. This might add a considerable processing delay.

Fuzzy MCDM algorithms are efficient to be employed when the handover decision includes imprecise parameters as well, like security or seamlessness and others, which might be evaluated using linguistic terms (high, low, very low, etc). Decision methods based on fuzzy logic alone are complex to use and require much expert knowledge and user involvement in order to make decision rules [CH92] In a preliminary phase of fuzzy MCDM calculations, the fuzzy data is converted into crisp numbers. If the fuzzy data are linguistic terms, they could be converted to fuzzy numbers using a conversion scale first. In a second phase, it is applied to classical MCDM methods to determine the rank of the alternatives. However, many of these fuzzy MCDM methods are cumbersome to employ, because fuzzy data are operationally difficult [Zha04].

2.5 Requirements and Challenges of Seamless Soft Handover

There are many efforts in the literature to address a set of requirements for any handover mechanism. The authors in [FK12] and [GK11] described system specifications and decision criteria for a sophisticated handover. Other than use case-specific requirements, the general specifications included a cross-layer processing of information, support of QoS, seamless mobility, support of intra- and inter-mobility, utilization of multiple wireless interfaces with caution to battery consumption, soft switching and security. Beyond RSS, a wide range of parameters are preferred to take part in the decision, for instance, user preferences, type of running application, service cost, network throughput, handover latency, and context information of the MN like velocity.

In this section, we briefly address a set of challenges for a handover system that includes all phases of a handover process. It assumes the availability of multi-homed devices and overlap of networks, and considers a reliable decision strategy that supports QoS in mobile communication scenarios.

For a soft handover, we define the following sub-tasks in accordance to the handover phases, and adopt them in our approach presented in the next chapters.

2.5.1 Gathering of Information

Parameters beyond received signal strength need to be considered. We categorize the interesting information based on the way of collecting it into:

- **Locally measured in the wireless interfaces:** Some of these are related to availability like RSS, while some are related to QoS like signal-to-noise ratio, delay and jitter.
- **Advertised:** these are provided by the network and are related to the offered services and network capabilities like cost and available shared bandwidth, or QoS classes if applicable.
- **User contributed:** the preferred network is one traditional input, but more user impact can be contributed, for instance by selecting one of the defined service profiles, which determines the weights of the criteria used in the decision algorithm.
- **MN context:** Speed and position are measurements that can be easily provided in a positioning-capable MN, to exclude non supporting networks for example. Consumed credit of data volume prior to throttling is also useful for HO decisions.

2.5.2 VHO Initiation

Here, we prefer to keep the MN looking for a better alternative even when not yet outside the coverage of a currently connected network. Each wireless interface in the MN can inquire a controlling entity for a permission to connect whenever its network type is in range. A VHO procedure might be triggered when an alternative network is found to be better than the connected one in terms of the defined criteria. Nevertheless, a stable status of connectivity is preferred, therefore, a mechanism that prevents switching back and forth between networks (ping pong) should be provided. For a soft handover, the concept of make-before-break should be implemented while being in coverage overlap area. If a handover is decided, the new wireless interface can connect before releasing the old one.

2.5.3 Decision Making

We want to select a network among a number of available networks with respect to different criteria, which is a case of an MCDM problem as mentioned earlier. However, increasing the complexity of the algorithm may add delay to the handover process, therefore, a robust but simple MCDM decision algorithm is needed. However, the management strategy should be independent of an algorithm's decision, and does control when a mathematical decision is needed, as described next.

2.5.4 Handover Management

To manage a VHO process, we need a centralized entity to control wireless interfaces, switch traffic in between and interact with the network layer entities to manage mobility. The controlling entity should provide the following:

- Consulting the decision algorithm, if the MN is in a state ready to check for a possible better network, where mechanisms should be implemented for stable handovers. Otherwise, the received interface inquiry is denied. Such decisions are not related to the mathematical algorithm, but rather to the management strategy.
- Advising the related interface to associate with the new network if a request is approved, and sending a disconnect message to the previous wireless interface when the traffic is already switched to the new selected one.
- This module should manage the VHO at L3 as well; modifying routing and interface tables and signal other network nodes to re-route traffic through an address resolution mechanism. As mentioned earlier, sessions might get broken when the IP address of a MN is changed upon a handover to a new network/subnet, unless some solution is conducted.

2.6 Discussion

As described earlier, soft handover and multiple criteria decisions are features of a sophisticated handover. Basically, we do not adopt information exchange systems that add signaling and introduce new processing load. We prefer an approach that resides inside the MN and makes use of the available information and events. Soft handover should be supported to utilize the co-existence of heterogeneous networks. For a fast process of decisions, we implement a simple but multi-criteria decision algorithm, namely MCDM-based algorithm. However, stable handovers should be considered in the handover strategy.

3 Mobility Management

Beside the actual switching in Handover, providing interoperability between heterogeneous networks and their addressing variants is very problematic [SD16]. The management of the node identity is, therefore, an essential part in the handover process. In this phase, the identification of the mobile node should be maintained regardless of its location, which is significantly important for running sessions when are desired to be preserved unbroken. This chapter describes addressing in IP-based networks, and mobility management in general along the layers of the TCP/IP model. The most recognized mobility approaches in the literature are discussed here.

3.1 IP-based Communication Networks

A very well-known term nowadays is *Internet*, which became an integral part of our daily life. It can be technically defined as the network of networks, which are referred to as subnets as well. Thousands of overlapping hierarchical networks interconnect to make up the Internet, however, the services provided over the Internet are due to various application servers deployed in many layers in the hierarchy [Sta09], [Tan02].

The Internet employs the TCP/IP protocol suite as a set of rules and protocols that regulates internetworking to enable communication between different networks and hosts. The Internet Protocol version 4 (IPv4) and version 6 (IPv6) within the TCP/IP are the main protocols used to route packets of data between source and destination hosts [For03] [Tan02].

The corresponding IP address version is assigned to subnets and hosts to designate their specific topological location. Conceptually, an IP address is a globally unique identifier of a host, and known as real/public IP address.

For an efficient use of the available addressing space and an easy localization, a kind of hierarchy is applied in the IP addressing scheme. Every host on the same subnet shares that subnet's address as a part of its IP address. Although schemes in IPv4 like Classless Inter-Domain Routing (CIDR) and subnetting keep a profitable addressing, if every host on every subnet had to have a public IP address, we would have run out of IP addresses years ago [Lam07].

A special set of IP addresses called private IP addresses are designed to save valuable IP address space and create a measure of well-needed security as well. It can be used on a private network, but they are not routable through the Internet. Hosts can have assigned private IP addresses, while sharing the same public IP address (gateway's IP address for instance) for communications on the Internet through NAT [SE01].

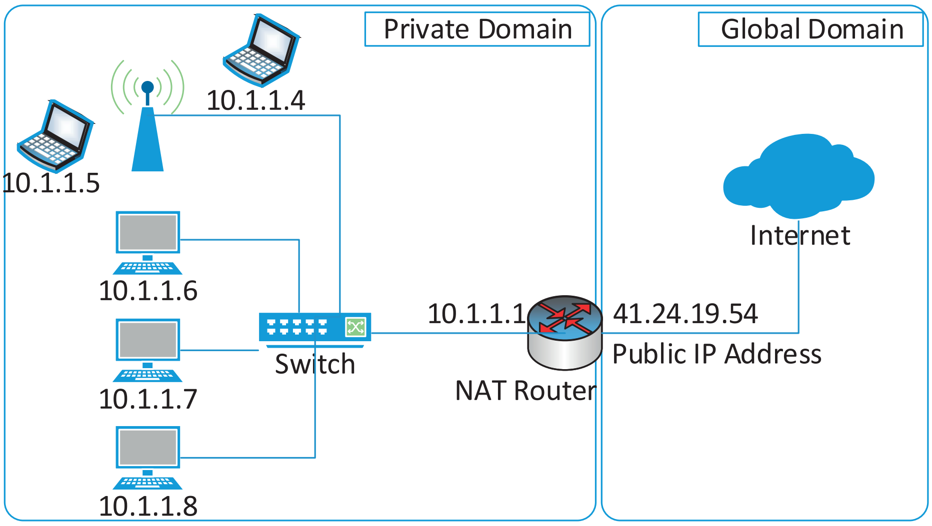

Figure 3.1: NAT between Private and Public IP address. The NAT Router maintains a cash to map and replace hosts actual IP private addresses and ports with the public IP address and also ports (generated by the NAT router).

NAT basically maintains a map to replace a host's private IP address with the global assigned one when packets pass through to/from the Internet. It uses Port Address

Translation (PAT) as well, where it replaces ports in the segment of the transport layer to distinguish between hosts when a packet is received from the Internet. Figure 3.1 illustrates a use case, where a company might have 10.1.1.x/24 private IP subnet address for hosts' internal communications, while a public one of 41.24.19.54/30 on the gateway interface to the Internet (/24 and /30 are the corresponding subnet masks).

IPv6 and its long address are introduced as the ultimate solution to the addressing shortage. However, for organizations that deeply invested in IPv4, it will not be easy to migrate to a totally new architecture and the process may take a long time.

Hosts can have benefits from the Internet using networks of wire and/or wireless access. In infrastructure-based wireless access, an AP provides the radio access to the wireless host into the network. A typical example is the connectivity through a WLAN. An AP might be also referred to as a Base Station (BS) in cellular networks generally, which were limited earlier to traditional voice services but started to present more data communications capabilities as well along its generations.

In wireless environments, mobility became one of the most common problems under research. A MN may move between different APs/BSs within the same or different networks. Switching connectivity between networks of different resources and services introduces, among others, identification problems, therefore, mobility can have a great impact on communication processes running on a moving wireless host.

QoS is a one major parameter in measuring the overall performance of wireless networks. Parameters like delay, jitter (delay variations), packet loss rate, packet error rate and bandwidth are the major metrics that determine the provided QoS [ssoI02, ssoI03].

There exist many challenges in maintaining connectivity for MNs and these are much more complicated if QoS and sessions continuity are desired to be preserved. Unfortunately, the TCP/IP model was not designed originally with mobility in mind. Many protocol extensions have been, and yet new solutions are to be, plugged into TCP/IP to support the continuous and various demands of the users and the evolved applications.

When thinking of mobility management, any solution should be able to provide a fixed addressing to MNs regardless of its location or the hosting network. This is particularly important for running sessions, where a MN is identified to other communications peers with its IP address prior to handover.

The cost applied to existing protocols and network's infrastructure in order to achieve mobility management should be, however, minimized in any proposed mobility solution. Other important parameters to consider also are in general security, flexibility and scalability.

3.2 Mobility Management along the TCP/IP Model

A considerable amount of work has been done in the field of mobility management since the presence of roaming and handover concepts, which can be addressed from different perspectives. The proposed approaches might be associated to one or more of the TCP/IP layers. We start first from the lower layers up to the application layer to give an insight on the most recognized solutions in this field.

3.2.1 Management in the Access Layer

One of the simplest forms of mobility management are these limited to the access layer. It might be referred to as layer 2 mobility/HO, where a MN moves out of the coverage area of its current AP/BS to another one within the same subnet (usually, inside the same operator network). Both APs/BSs belong to the same network in this type and a MN might preserve its IP address, therefore, no identity issue arises in such types. The main task in such type of mobility is to search for an adequate AP/BS to re-associate with. However, if the newly selected AP/BS belongs to another network, a higher layer should be involved in the process.

A handover into a new network means the assignment of a new IP address to the MN. It is, therefore, referred to as layer 3 mobility, where the identity (IP address) by which a MN is reachable is affected. For such a handover, approaches should be implemented in layer 3 or in the layers above as well, with no specification to a certain technology in the access layer.

3.2.2 Management in the IP Layer

Implementing solutions at layer 3 of the TCP/IP model usually includes changes in the network infrastructure. A route to the MN should be achieved, which can be realized by the employment of an indirect node, mostly known as an agent, which forwards the traffic from its primary path to the visited network [Edd04]. Well-known layer 3-based mobility protocols are Mobile IPv4 (MIPv4) and Mobile IPv6 (MIPv6) standards for the two corresponding versions of the IP protocol [Per02, JPA04]. In next, we briefly introduces these protocols and some of their extensions.

Mobile IP

The basic idea in mobile IP is to achieve a transparent routing of IP packets by tunneling, independently of the MN location. Basically, tunneling is a process of encapsulating a network layer protocol inside IP packets to be delivered to the network in corresponds to the inner packets. Considering MIPv4, a tunnel is established between a node called Home Agent (HA) inside the *home network*, to which the MN belongs, and a node called Foreign Agent (FA), which the MN has moved to. In this way, IP packets destined to the MN using its old IP address are embedded inside other IP packets that contain the actual address of the visited network. There, a Care of Address (CoA) represents the point of attachment of the MN and the end point of the tunnel. This address usually represents the FA's interface on the foreign network being visited by the MN. Nevertheless, a MN keeps using a longterm IP address known as the Home Address (HoA).

Upon a handover, the HA maps the HoA and CoA registered by the MN, which may register it also with its communication peer, known as the Correspondent Node (CN)Routing through the HA introduces the *triangular routing* problem as can bee seen in figure 3.2. However, route optimization might follow this procedure to enable direct communications between the MN and the CN, if the gateway in the visited network allows it. For this, the CN has to cache the MN binding to enable direct tunneling to it.

Due to the use of a non optimal route through the home agent and the tunnel establishment procedure, a considerable delay by extra signaling is introduced, which is harmful for session continuity during handover. However, the CN requires modifications to support the mobility protocol when route optimization is to be carried out, which is anyhow, applicable only after a successful handover and has no effect on reducing the HO delay.

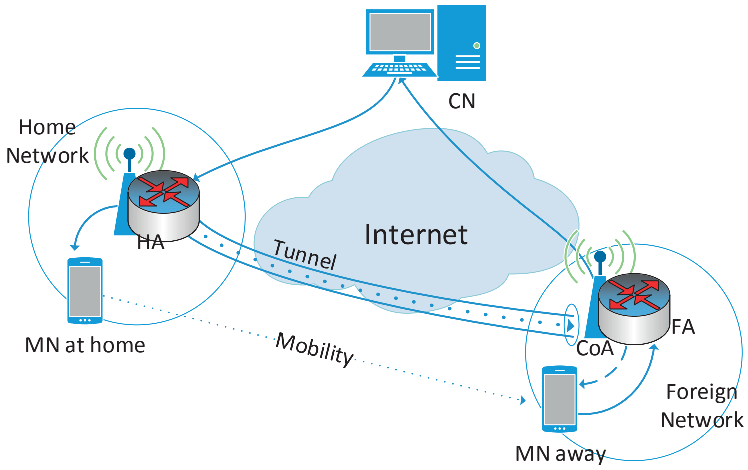

Figure 3.2: Basic Mobile IP. The traffic destined to the MN passes through HA, which forwards it to the registered CoA.

MIPv6 is similar to MIPv4. No FA is used in this version, where the MN represents the other end of the tunnel in the foreign network. The MN obtains a new IPv6 address, which represents the CoA, in the new network. It informs the HA and later the CN about its location using Binding Update (BU) messages. Among other differences in the signaling scheme, MIPv6 comes with improved security features. However, the same disadvantages of MIPv4 apply also in MIPv6.

Extensions like Fast MIPv6 (FMIPv6) [YCK$^+$10], Hierarchical MIPv6 (HMIPv6) [SCEB08] and Proxy MIPv6 (PMIPv6) [GLD$^+$08] have been proposed, which extend the basic MIPv6 to overcome its disadvantages.

Mobile IP Fast Authentication (MIFA) presented in [Dia10] assumes mobility to a limited number of pre-defined neighbor network. Sharing MN related information with these networks fastens authentication process and thus, supports a quick session continuity upon handover. However, the basic assumption of pre-knowledge and co-operation of neighbor networks might represent a restriction. Also MIFA outperforms MIP only at high velocities.

Other than the MIP concept, the Location Identifier Separation Protocol (LISP) [FFML13], the Host Identity Protocol (HIP) [MNJH08, NGH10], the Location Independent Addressing for IPv6 (LIN6) [KTI03] and the Mobile NAT (MobileNAT) [BHSM05] implement a new semantic for addressing, which will be described in this section, too.

These extensions and enhanced approaches will be explained in more details in the next sections.

Proxy MIPv6

PMIPv6 is a network-controlled management protocol. In its architecture, there are two main entities; a node called Mobile Access Gateway (MAG), which is responsible for detecting the MN mobility and its association to the APs, and a node referred to as Local Mobility Anchor (LMA), which is similar to the HA in MIPv6. Acting as a proxy, the MAG of a visited network does the MIPv6 signaling on behalf of the MN. It handles also binding updates, tunneling and traffic forwarding to the MN. The main advantage in this protocol is the elimination of the risk of dropping MIPv6 signaling messages on the wireless connection during handover. Figure 3.3 shows a simple topology for the mentioned protocol.

Hierarchical MIPv6

HMIPv6 is an interesting enhancement to MIPv6. It enables less signaling in cases of local mobility. It introduces a node called Mobility Anchor Point (MAP) inside the visited network, where it acts as a local HA to the MN. A MAP may be connected to

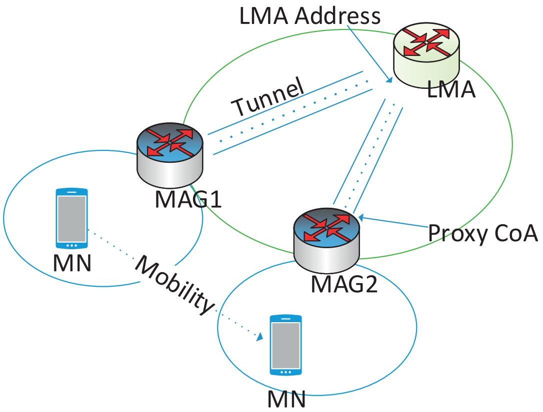

Figure 3.3: Proxy MIPv6. The MAGs undertake the MIPv6 signaling on behalf of the MN.

more than one Access Router (AR), as can be seen in figure 3.4.

A MN receives a Regional CoA (RCoA), which is related to the MAP that registers the MN, and a Local CoA (LCoA), which is related to the ARs. RCoA is the address registered by the HA as the MN's CoA. One or more MAPs can be deployed in domains and in a hierarchical way, however, a MN must register to only one MAP. When a MN moves within the domain of its MAP where it has registered, it updates only the LCoA to that MAP. The signaling runs in this case only locally without having to notify the HA. Whenever a MN roams into a new domain, it will be represented with a new RCoA from the new MAP and eventually, obtain a new LCoA from an AR. In this case, which is considered as a global mobility scenario, BU should be sent to the HA to update the new RCoA (known as CoA to the HA).

The described scheme reduces unnecessary signaling to the HA in cases of local handover. Still a crucial factor is the selection of the MAP, to which a MN should register when visiting a new domain. For instance, registering a MN that has frequent handovers to a low MAP in the hierarchy would result in many BUs to the HA, and thus falling back to the basic MIPv6 behavior. However, this approach still suffers from lengthy handover and packet loss with the global HO, which is a scenario we can not assume that it would not occur so frequently, especially with high speed

[KM15, LHM10].

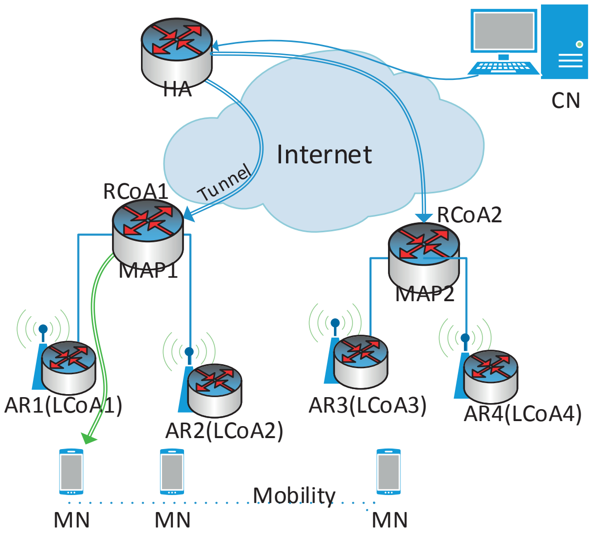

Figure 3.4: Hierarchical MIPv6. HA knows only about the RLoCs, while MAPs handle local mobility when the LCoA is changed.

Location/ID Separation Protocol

LISP is a protocol aiming at optimizing global routing. It splits the device identity by introducing two numbering spaces; the End point Identifier (EID) that is not routable globally, and a routable Routing Locator (RLOC) [FFML13]. RLOCs are IPv4 or IPv6 that are used for routing through the transit networks.

Mapping between the two should be maintained in a mapping sub-system connected to the Internet. Host's applications bind to host's EID, which is used as the address for the transport layer connections. Once the RLOC associated to an EID is resolved, packets with EID addresses are encapsulated in a second header of RLOC addresses. Special routers called tunnel routers are used for encapsulation and decapsulation purpose at each LISP-enabled network. In figure 3.5, PC1 knows only the EID of PC3. It is up to Router1/Proxy1 to inquiry EID-to-RLOC resolution and tunnel traffic towards Router2/Proxy 2 in order to enable communications between PC1 and PC3.

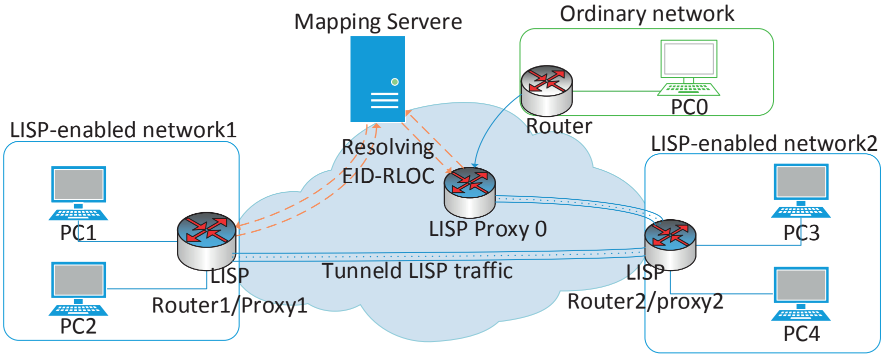

Figure 3.5: LISP typical communication.

In [FFLM14], a support for mobility is added (see figure 3.6). The mapping is maintained and updated upon each handover and attachment of the MN to a new network, where it receives a new RLOC. The EID-to-RLOC mapping should be updated also in the cache of all the communication peers. A tunneled communication between a proxy router and the MN should keep a fixed identification of a MN [NJP+13]. This protocol avoids traditional HA-based solution (in MIP), nevertheless communication with non-LISP networks requires the use of a proxy router.

However, the long headers used by this protocol can increase the Message Transfer Unit (MTU) size, which is problematic in tunneling approaches, and the mapping update mechanism can introduce delay in re-routing sessions to the new location.

The aforementioned HIP and LIN6 are also based on the idea of introducing two identities; one fixed for sockets associations, and one (IP address) for topological locations in the network. Similarly, updating the mapping sub-system might introduce delay to traffic recovery in case of a handover.

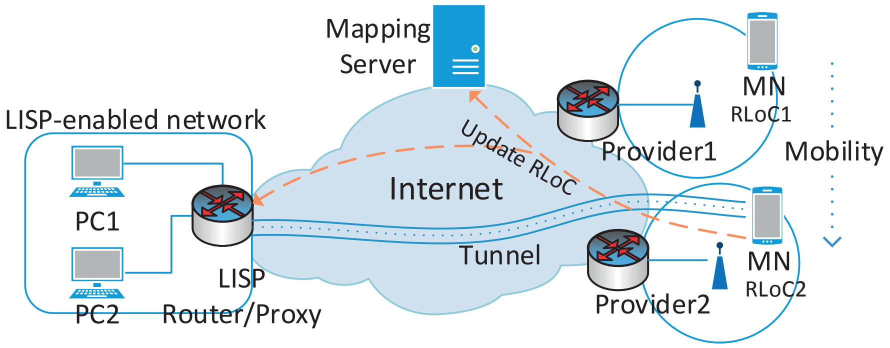

Figure 3.6: LISP Mobile Node. A MN is the LISP site in this case, where encapsulation and decapsulation is achieved inside the MN.

MobileNAT

An approach to manage mobility using NAT is presented in [BHSM05] and named as MobileNAT. It uses two IP addresses per MN; one called A_v for identification at the application layer, and a second called A_p for actual routing at the network layer. They both can be private or public address. A_v is a fixed identifier, while A_p represent the current point of attachment and hence, changes due to mobility. It has routing significance only within a domain and therefore, can be a private IP address. Since A_v is not used for routing, the MN must translate it to its A_p to make packets routable before sending them to the Anchor Node (AN) in the domain. A thin software layer, called *shimlayer*, is placed above the access layer and used to maintain translation rules inside the MN.

A modified DHCP is employed to supply the mentioned addresses for hosts and to signal mobility events to a signaling node called Mobility Manager (MM). The MM is used to signal the changes in packet processing rules to the node that performs the NAT in the domain, usually the AN, in the event of node mobility. When a MN obtains new IP addresses (A_p,A_v) due to mobility, the DHCP conveys the event to the MM, which in turn signal the changes in the mapping rules to the AN (see figure 3.7). This represent an intra-domain case and results in directing the traffic by the AN to the new subnet.

When the MN moves to a new domain, which is a case of inter-domain mobility, the previous serving AN, refereed to as Home_AN, must tunnel the running traffic and forward it to the new AN, which referred to as Visited_AN. This, however, represents a non-optimal route for running sessions, similar to mobile IP scenarios.

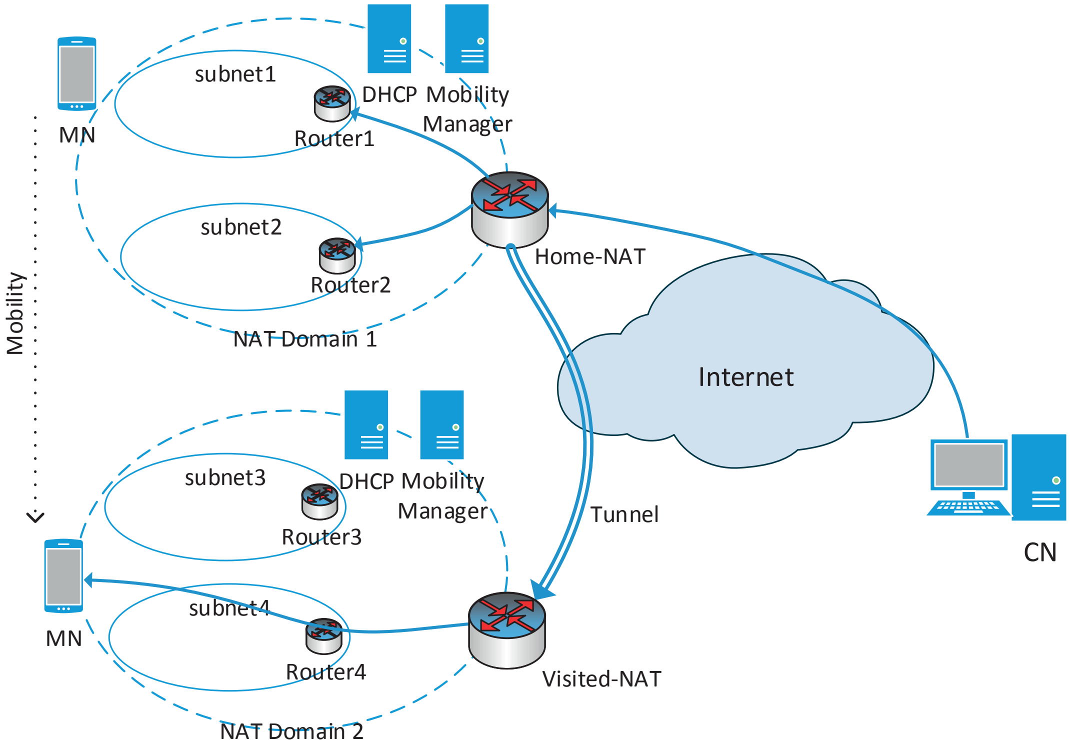

Figure 3.7: MobileNAT within intra- and inter-domain mobility scenarios.

A new generated traffic will use the new acquired A_v, while those started in the home network must remain using the old A_v. This might results in an address aliasing problem related to duplicated A_v addresses. It can happen that another node in the domain is already using the same A_v address assign to a MN from a previous network. Such aliasing prevents the communication between the MN and any other node until old sessions are over and the previous A_v is deleted. Another solution is to avoid address conflicts by administrating non-overlapping ranges for the A_v IP address. This represents after all, a strict constraint, which is useful only within a single provider network of multiple NAT domains.

3.2.3 Management in the Transport Layer

A solution implemented at the transport layer is considered as an end-to-end approach, where it is up to the end points, the MN and its communication peer, to manage the mobility, without changing the network infrastructure. This requires in principle, however, updating the communication peers upon each handover.

Mobile Stream Control Transmission Protocol (MSCTP) extends the basic SCTP to support handover and multihoming [TR05]. It maintains a list of IP addresses of the MN and modifies it in the SCTP association through the exchange of configuration messages. Upon a handover, a MN notifies the CN regarding its newly obtained IP address to add it to the list of the connection's IP addresses, where the old address can be then removed. Mobile Multimedia Streaming Protocol (MMSP) [MYO03] assumes overlap in the coverage of two networks, where a second interface can associate and obtain a new IP address. The CN maintains a list of destination addresses for the MN and will add this new IP address to redundantly send packets during handover. Upon a completion of a handover, the IP address of the first interface is removed. To validate such approaches, high cost is applied by the necessity to modify each single potential communication party in the Internet to support mobility.

MSOCKS introduced in [MB98] also supports multihoming and mobility, but only locally. It uses a proxy server in the network to be a point in the middle for the MN communications with the CN. It modifies the basic SOCKS by identifying MNs-proxy sessions and adding a signaling messages to replace the MN IP address in this part of the connection in case of a handover. The proxy-CN part of the communication remains unaware of the change. However, the assumption of only local mobility scenarios can be seen as a restriction.

Another solution deployed nowadays is multi-path TCP (MPTCP) [FRHB13]. It creates TCP sub-flows for each available interface to maximize the utilization of resources, which is in the same time useful for mobility since the creation/disappearance of sub flows is transparent to the application. This solution suffers from middle-boxes in the Internet, where changes to any sub-flow's header information breaks it, and hence the protocol falls back to the basic TCP [RPB$^+$12]. Nevertheless, the CN should support the protocol also. [IS15] describes an experimental protocol atop of UDP and optimized for HTTP/2 semantics. Limitations to specific applications, the need for modification at the CN also, and the lack of measurements for analysis apply restrictions in the deployment for further research.

Roaming-Enabled ArCHitecture (REACH) [ES08] employs a proxy server as an anchor point for MN's communications. An entity called REACH-client resides in the MN and is responsible of intercepting data packets from the application layer using relay plugins. The intercepted data are sent over stream-based or datagram-based logical links for UDP and TCP traffic respectively to an entity called REACH-server in the proxy server. The proxy server forwards data to their destinations. Although handover breaks transport associations between the REACH client and server, the logical links remains unaffected, where the associations are designed to be short-lived and are reestablished frequently. The proxy server installation might introduce, however, non-optimal routes in case of mobility.

3.2.4 Management in the Application Layer

An application-based solution extends the IP telephony infrastructure to handle mobility. Like some former class of approaches, it does not apply changes to the existing network infrastructure, but requires updating the communicating nodes upon each handover. It extends the Session Initiation Protocol (SIP) [RSC$^+$02] for this purpose. SIP binds the user identifier to its IP address. In case of a handover, a users registers the newly assigned IP address with the home SIP server, which will reply to new sessions' requests with registered new MN's location information. If handover occurs while a session is running, the MN invites its communication party to reestablish the call using the same session's identifiers, and provides it with its new location information in the same time. SIP-based approaches are suitable for IP telephony, which are UDP-based applications [WS99] and requires the other communication party to support SIP, too. SIP is, however, well known that it introduces intolerable delays [FK12].

3.2.5 Management in Multiple Layers

There exist many other approaches that can be related to more than one layer of the TCP/IP model and are, therefore, classified as cross-layer solutions. The idea is to optimize mobility management through the use of information from more than one layer. For instance, some QoS-aware handover approaches conduct resource reservation when additional information are supplied from the application layer [GSRM05]. In [PCT03], information regarding running applications assist in triggering one of two combined mobility solutions like MIP for TCP applications and SIP for real-time. Similarly, link layer can state useful information of interfaces' parameters to the employed management solution.

Fast MIPv6 [Koo05] extends MIPv6 to introduce less handover delay. It makes use of the discovered list of APs in the link layer to request subnet prefix information from one or more APs from this list in prior to handover. Once the MN switches to the network of one of these APs, it retrieve the related information from the cashed AP-prefix mapping to reduce configuration time. BU, tunnel establishment and non-optimal routing are still, however, issues inherited from the basic MIPv6.

QoS-aware Mobile IP Fast Authentication Protocol (QoMIFA) is a hybrid approach presented in [Aln12]. It combines QoS Resource Reservation Protocol (RSVP) with MIFA [Dia10] to provide guaranteed QoS upon handover. It introduces a new entity called mobility object to RSVP in order to encapsulate MIFA control messages. The same limitations of MIFA apply here, though exploiting information from other layers might be useful to enhance the handover.

Approach	**Concept**	**Advantages**	**Disadvantages**
MIP	tunneling, BU	transparent to higher layers	non-optimal routes and high latency
PMIPv6	tunneling, proxy-based	signaling on behalf of MNs	non-optimal routes, and only network-controlled
HMIPv6	tunneling, hierarchical setup	reduced signaling	requires cooperation of networks, basic MIP behavior in global HO
MIFA	MIP-based, pre-sharing of MN's information	outperforms MIP in high speed	has MIP disadvantages, restricted assumption
LISP	two identities, mapping	avoids MIP concept	requires proxies for non-LISP sites, delay due to update of peers
HIP and LIN6	two identities	transparent to upper layers	delay due to update of peers, modifications cost
MobileNAT	two addresses for MNs, mapping	transparent to upper layers	non-optimal route in global mobility
MSCTP and MMSP	a list of MN IP addresses	supports multi-homing and soft HO	modifications cost
MSCOCKS	proxy-based	CN is unaware	only local mobility
MPTCP	sub TCP flows	max. utilization of resources	suffers from middle-boxes
REACH	proxy-based	mobility hidden to CN, supports TCP and UDP	non-optimal route
SIP	SIP server as a HA	direct path between MN and CN	CN must support SIP, only for UDP apps., and delay
FHMIPv6	pre-request for prefix info. using L2 info.	reduced configuration time	non-optimal route
QoMIFA	MIFA-based	QoS guarantee	MIFA restrictions

Table 3.1: Comparison of the described solutions for layer 3 mobility

3.3 Discussion

As a summary, table 3.1 shows the described mobility solutions with a brief keynotes on their concept, advantages and disadvantages. From the described approaches, we have addressed the main features for the management mobility approach to be presented in this dissertation. Our idea is to keep the issues of mobility as close to the MN as possible and, hence, not to apply any modifications on the CN's side, because modifying servers that are already installed and running all around the world means a huge amount of efforts, cost and risk. Modifications inside TCP/IP are also important to be minimized, so we do not adopt a totally new naming space to maintain fixed transport associations, but rather make use of already-available schemes, as will be described in the next chapters. Therefore, we adopt primarily a network layer-based management solution. Additional information from other layers are useful to optimize the process and, therefore, we address a network layer-based hybrid approach that makes use of lower layer information, and input from application layer as well. However, the point in the middle-problem had to be taken into account to avoid or minimize the effect of non-optimal routing issue.

4 Proposed System for Vertical Handovers

This chapter gives an insight into the proposed solution and the modules implemented in the MN, and other nodes in the networks as well, to validate the suggested system. It describes the sub-tasks handled by each node within a typical communication scenario while having a vertical handover.

According to the reviewed phases, the challenges, and the requirements mentioned earlier in section 2.5, we present our cross-layer system and framework inside a MN to achieve the tasks required for soft, seamless and QoS-supportive vertical handover. The framework consists of three modules namely, the Controller to enable soft handover, the Decider to compare the available alternatives, and the DiNAT-agent for address resolution purpose upon handover and the assignment of a new IP address to the MN. The description of the proposed system and its modules, including simulations and validation results, have been published in [ARS15, ARS16, ARAS16b].

4.1 Information Gathering and Initiating a Soft VHO

The Controller represents the central module that manages the VHO process in the proposed solution. It acts as a controlling and overlay entity to the wireless interfaces available underneath in the access layer (see figure 4.1) regardless of its radio technology. The two other shown modules are inquired to act by the Controller according to the phase in process, and as will be described next. The Controller receives requests for permission-to-associate from each wireless interfaces whenever being under the coverage of the network type related to that interface.

An important note here is that the process of horizontal handover management is separated from that of vertical handover in this system. The selection among access points of a certain network type is a task of the wireless interface and its corresponding entities. Nevertheless, the proposed model can be easily modified to manage horizontal handover as well, but following the strategy set for the vertical type.

Soft handover is enabled through the concept of make-before-break. It assumes the availability of multi-wireless interfaces on MNs and an area of overlapped coverage, where an old connection will be disconnected only when the new one is definitely associated, configured and ready to process traffic, as figure 4.2 illustrates. In this sequence chart, a node (MN) starts communications with a server (App. server), first utilizing network (Network_A) and then switches to network (Network _B), where both belong to the same domain and have access to the Internet through the node (Gateway).

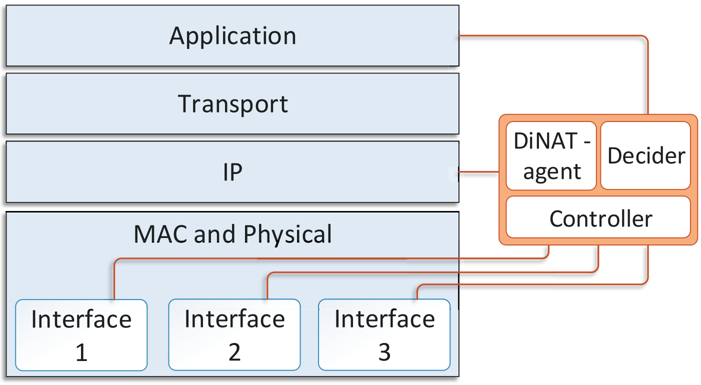

Figure 4.1: The VHO cross-Layer framework along the TCP/IP suite model

The Controller applies a handover management policy of connectivity, but inquiries the Decider to compare between the available alternatives. As a part of the connectivity policy, it configures a set of timers, mostly for stable connectivity, that define its state. The decision algorithm is inquired for a decision only when the Controller has no any constrains against a change of connectivity.

According to its state, the controller decides whether to consult the decision algorithm or to directly reject a request with a deny message. Parameters of node velocity and user's best ever preferred network play also a key role in such direct decisions that might be taken, without the need for the mathematical decision algorithm.

Other parameters regarding the handover policy, like for example, excluding specific networks due commercial agreements, might be added to the policy applied by the Controller module.

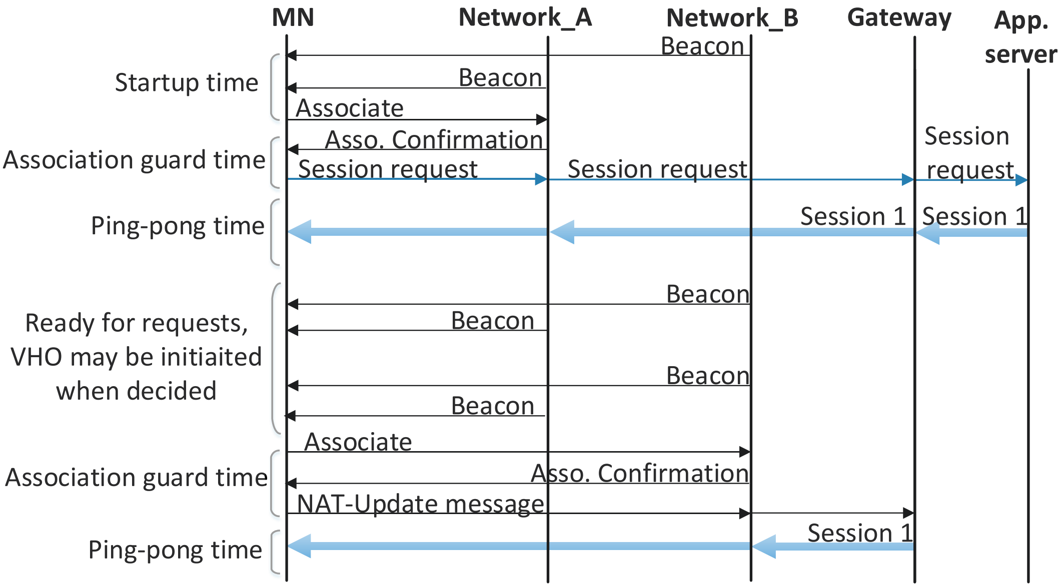

Figure 4.2: Network selection and soft handover procedure. Inter-mobility is shown in this scenario. The AP of each network is omitted for simplicity

The timers and the related Controller's states are:

- Initial request timer: It starts when receiving the very first *request-to-associate* from an interface, which takes place usually when switching on the MN. The goal is to fairly discover all the networks in the surrounding. This timer should be configured with a sufficient time to scan the channels of the available networks. The Controller is in the state *Waiting Startup* during this time, and changes to state *Association guard* when the time is out, where the best request is granted the connectivity, while the others are denied. In case the time is out when a MN is switched on in an area with no coverage, the Controller changes to state *Ready for Reqs*, in which the Controller will inquire the decision algorithm whenever any request is received.

- Association guard timer: This runs when a request has just been approved, to wait for the association procedure to finish before setting the related wireless interface as the active one and switching traffic onto it. A timeout event leads to state *Ready for Reqs*, while a successful association stops the timer and leads to state *Pingpong guard*.
- Pingpong guard timer: after a successful VHO, requests within this time will be rejected to prevent back and forth switching. When this timer ticks off, the controller will be ready again to examine new requests through its decision algorithm while being in state *Ready for Reqs* again.
- Dual mode timer: After switching to the new wireless interface, the old one can still receive packets (some might be still on the air) during this time. This timer has no effect on the states of the Controller.

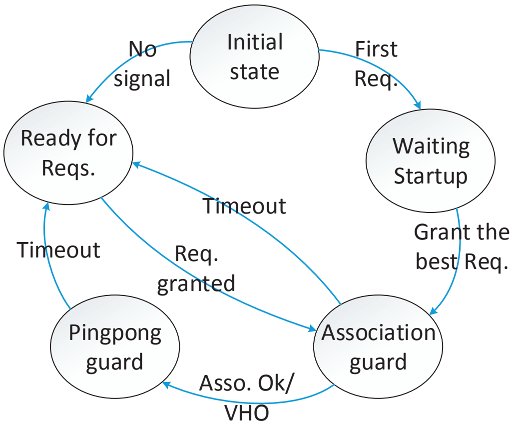

Figure 4.3: The states of the VHO Controller

Figure 4.3 illustrates the different states of the VHO controller. If the controller is in a state that allows to examine requests (*Waiting Startup* or *Ready for Reqs*), the decision algorithm will then decide whether a candidate network is a choice better than the previously selected one. When a request (an alternative network) worths a handover, the Controller grants a *permission-to-connect* to the corresponding interface. It waits until a confirmation of association (event *Asso. Ok*) has arrived to act then at the network layer by triggering the address resolution procedure.

4.2 Decision Making

The Decider module implements a configured decision algorithm to compare between the available networks according to the defined set of decision criteria. This comparison is frequently performed between the current connection and any other available one. It could be also carried out to choose between many alternatives when a MN has lost connectivity or has just been switched on and is trying to acquire connectivity to a network.

A decision algorithm compares here between only two network at time. The reason is to carry out the comparison while being on the move, rather than creating a list of alternatives and waiting for an amount of time (to receive beacons of all network), during which networks conditions might change due to mobility. If a plenty of network are available, each will be compared to the currently connected one in a serial form. It is up to the timer set in the Controller to prevent unstable switching, as mentioned earlier.

The Decider plays no role in the handover policy/strategy, which is handled by the Controller, but it runs a set of mathematical analysis and calculations to decide which alternative is the best according to the defined set of criteria. It receives decision inquiries from the Controller, which holds the attributes of the two networks it is trying to evaluate. The Decider returns a result to the Controller, which in turn is responsible of triggering the subsequent events in case a handover was decided. The result of a comparison is measured up to a configurable offset to decide whether it does worth to declare one network as better than the other.

The decision algorithm could be a traditional RSS-based one. However, when a handover is desired to be sophisticated with QoS support and more user input to the decision making, an MCDM algorithm is to be employed.

For flexibility, a configuration file in the simulation program can activate the desired algorithm according to the user selection. The considered decision criteria are the RSS, the SNIR, the available data rate of the access point/network, the service monetary cost and the delay and jitter between a MN and the network access point.

RSS and SNIR values are normalized to enable comparisons between systems that have different power constraints. Each is calculated as the ratio of the actual measured value to its threshold values. A MN might have a usage profile with each of the networks, which may allow a high bit rate until some consumption limit of data volume, after which it gets a low bit rate. Therefore, the bit rate value considers not only the available rate of the networks, but also throttling (to some configurable offset), when applicable, to calculate the real bit rate after a possible near reduction, if the user already consumed a big amount of data within the examined network.

The criteria weights are set in pairwise matrix as mentioned. Even so, it is undesirable to bother a user of a smart phone with many detailed and different sets of criteria weighting. Therefore, the idea is to have it configurable in details only for research work but, to be packed as few service profiles able to be activated by the user to apply a desired set of weights to the decision algorithm, for example, to promote cost effective, highest quality services or any other usage plan. Nevertheless, a user is still free to bypass the decision algorithm and specify his preferred network type.

4.3 DiNAT for Mobility Management

The DiNAT-agent module handles the identification issue of a MN when switching to a new network and hence, being assigned a new IP address. As the name implies, it employs the NAT function but extended to enable a dynamic update of its entries to support mobility.

The Basic Idea

In almost all of the networks and due to the lack in the addressing space of IPv4, NAT and PAT are used to map local IP address to global ones.

IPv6 has no shortage in addressing, but NAT might still be used for purpose of smooth transition to the new IP version, and it is useful when migrating big networks to a new Internet service provider for instance, to eliminate the need to renumber each individual machine on the network. Such NAT types are called IPv6-to-IPv4 NAT (NAT64) [BMvB11] and IPv6-to-IPv6 NAT (NAT66) [WB11] respectively.

This widely deployed technique can be employed to support mobility and handover. To accomplish this, we leverage NAT and spread it on levels for purpose of path segmentation. A major component of the approach is a proxy server/node connected to the Internet to apply NAT, however, between public IP addresses as well. The idea is to hide the changes of the addresses in the networks from communication peers in the Internet and hence, keeping a fixed identification of a MN to them.

Nevertheless, NAT of private to global IP addresses at an operator gateway would not break the traffic but rather can further be utilized to support intra-domain mobility. A NAT server in the proposed system is able to dynamically update its NAT entries upon receiving an update message from the mobile device in case of a handover. We name this approach DiNAT, where entries are pron to MNs dynamic changes in addresses, and also to differentiate it from a NAT type known as dynamic that shares a single public IP address for many private ones. DiNAT focuses on IPv4 scenarios, but as mentioned, is applicable in concept for IPv6 as well.

DiNAT servers are deployed in two layers; local and global. Locally, NAT functionality is already used in gateways of most of the IPv4-based networks to map global IP addresses to private ones. We leverage this basic function to support intra-mobility. Furthermore, a cooperation with another DiNAT proxy connected in the Internet (globally), which performs a NAT between global addresses, is able to support global/inter-mobility. We assume the availability of a Network (or domain) Identifier (NID) to recognize domains of mobility. It is useful to deploy several DiNAT servers in the global network to spread load and for route optimization purpose. A MN is not permanently attached to a DiNAT proxy node but rather, selects a new one based on routing metrics and load parameters suitable for each new session to be started, after HO into a new network for example. A signaling scheme upon the start of a session runs to identify the potential DiNAT proxies and select one. Figure 4.4 shows the described topology.

The availability information of proxies are assumed to be supplied by the network operator/service provider, if it supports the solution, otherwise the MN uses the information provided by a previous supporting one. The selection criteria are the routing metrics of a proxy node (towards the MN and the CN) and its process load. This releases the MN from being attached to a fixed node, like the HA as in MIP. The DiNAT proxy node updates its mapping in case of HO. A MN might be communicating through multiple proxy nodes for multiple sessions that started in different locations.

4.4 Handover Operation Overview

An insight of operation is provided here involving a communication session between a MN and an application server in the global network. Intra and inter domain mobility cases are described here as supported by the proposed scheme. We denote the aforementioned global and local servers with Global Translation Server (GTS) and Local Translation Server (LTS) respectively. The traffic flows through an LTS, which represents the gateway of a DiNAT-enabled network, and a GTS, which resides in the global networks. A specific GTS is selected among others through a selection mechanism triggered by the MN. The employment of an LTS and GTS together applies a hierarchy of two levels, where the lower (LTS level) supports intra-domain mobility and the higher (GTS level) supports the inter-domain mobility. However, GTS can be used alone when a provider's network is not aware of the DiNAT implemented approach.

The MN gets assigned a private IP address using DHCP, which is a very common scenario in IPv4-based network. We assume that the MN is the node that starts a session with a server residing in the Internet. However, if the MN is acting as server inside a private domain, port forwarding scheme (it uses the gateway's IP address for public advertisement in this case) or static NAT (one-to-one mapping of private and public IP addresses) could be used to make the CN reachable from the global domain. In this case, the relevant DNS server should be updated to get the MN reachable with the new address upon a handover. For a running session, the same DiNAT update procedure is applicable conceptually. IP address translation inside the MN (between the previous IP address and the new one) is then useful to hide the changes from the transport association.

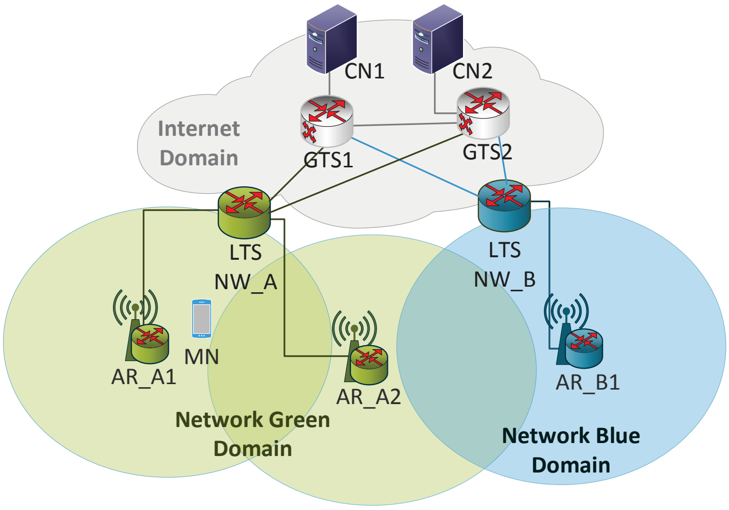

Figure 4.4: Network topology

4.4.1 Signaling

In this section, we provide a description of the signaling that takes place in the MN, at the network layer to establishing a communication session, and at the link layer while mobility.

4.4.1.1 Signaling in the Mobile Node

At the lower layers, the MN associates to one of the available networks based on decisions of the Decider module. Part of the parameters delivered to the Decider are values advertised by the APs through beacons, like available bandwidth, cost of service. These, however, could be interesting to the user in HO decisions.

To be always best connected, the Controller continuously receives requests to associate from the interfaces that have their corresponding network types in coverage as mentioned earlier. Whenever a better alternative is available (to some offset), the Controller switches the stream to the newly selected interface after passing the timers necessary for stable handover. Other than the update procedure prior to handover, the MN receives some extra information when exchanging DHCP messages to fasten resume of traffic in the new location. The MAC address of the gateway and the NID are also supplied besides the IP information. This saves the time consumed by the ARP in the MN when starting communication.

At the higher layers, the DiNAT-agent starts a procedure to set and select a GTS server for the session to be started. This is triggered when the application layer sends a Session Request (SR) to start one. The network layer retains this message until the DiNAT-agent allows to release the SR, when the GTS selection procedure is completed.

To choose a GTS, the DiNAT-agent contacts a group of nearly available GTSs. The knowledge of GTSs is made available by the LTS, which can find a relevant group by the use of an any-cast scheme for instance, or have them statically configured by the network administrator.

The MN contacts the provided set of GTSs to query performance parameters using a message named Server Metrics (SMT). It contains fields of:

- Message identity: to identify the request/reply in this MN-CN association.
- Message type: whether it is a request or a reply of information.
- CN's IP address.

- Metrics: the metrics in the routing table of the node, both towards the MN and the CN as well.
- Load: the processing load of the node.

This request-reply paradigm to acquire information is implemented over UDP for simplicity, where we designate special ports for this purpose, anyhow, it can be also carried out using other basic signaling applications like HTTP [FGM+99]. Figure 4.5 in the next section shows a signaling denoted as *Req for Info* and *Info Reply*, which represents an exchange of this message to select a GTS in prior to a session request.

The DiNAT-agent selects one of the defined GTSs using any simple selection algorithm. The one implemented in our framework compares the node's load first to select the lowest and then the routing metrics to select the best.

When a GTS is elected, the SR is sent out of the MN. The GTS forwards the SR and sends back a message through the LTS called Session Creation Information (SEI), which conveys MN and session identification information. The MN uses the SEI to fill the relevant entry inside a repository maintained locally inside the MN called Sessions Table that contains identifications for the sessions and the coupling of the MN-GTS-CN nodes. The session table contains the following field:

- Protocol: to identify the protocol used in the transport layer.
- MN Port: the source port number used by the transport protocol.
- CN Port: the destination port number.
- CN Addresses: the IP address of the CN.
- MN Address: its IP address at the time of the session start. This enables supporting continuity for connection-oriented application as well.
- GTS Address: the IP address of the GTS selected for this session.
- Global Session ID: an index to identify the session within this specific MN-GTS-CN association.

The last two are determined only after the receipt of the SEI from the GTS, while the other are set locally by the MN at the start of the session.

4.4.1.2 Signaling in the DiNAT Servers

Beside providing ordinary DHCP service to MNs, an LTS provides also information related to the available list of GTS servers as well. It also forwards SMT messages between a MN and the available GTSs to facilitate the selection of a GTS to a communication session.

Upon the receipt of the SR message sent by a MN to start an application, it creates the corresponding NAT entry for this message and forwards it to its destination, which is the selected GTS. This step is confirmed to the MN by an SEI message, which identifies the session entry in the LTS and the GTS as well. A unified index number is shared between the MN, the LTS and the GTS to validate the proposed approach. Together with the IP address information and port numbers, it identifies the flows of traffic within the three mentioned nodes.

A GTS exchanges SMT messages with the MN when a communication session is to be placed. Later on, they also exchange messages related to the update scheme within the suggested approach in cases of handover.

4.4.2 Traffic Flow

This section gives an insight into how traffic flows are treated between the MN and a communication peer in the Internet. Figures 4.4 and 4.5 can be followed to better understand the description of the DiNAT behavior.

4.4.2.1 Uplink Stream

When a MN wants to start a communication session with some other node in the Internet, an application server for example, it sends first a request to start the session.

This messages helps to setup the flow NAT entries in the intermediate relevant nodes after a successful selection of a GTS. The packets in the uplink direction are tunneled towards the GTS. The outer packets are destined to the GTS while the inner has the CN as the destination. Tunneling is necessary here to enforce the packets to have two destinations, the GTS as an intermediate node and the communication peer.

While passing through the LTS, NAT is applied to the tunneled packets (to outer addresses). These, however, are untunneled in the GTS, applied to NAT and forwarded to the CN. If an LTS or a GTS has no entries for a packet, it creates one and forwards it to its destination (relevant to the first packet sent). All traffic arriving the CN has a source address of the GTS's IP address. Therefore, to the CN, its communication peer is the GTS. A GTS also sees only the LTS IP address in the source address of packets sent by the MN, and only the LTS faces the actual address of the MN.

4.4.2.2 Downlink Stream

The packets sent by the CN are destined to the GTS, which searches its NAT table to forward them to the corresponding LTS. The later also applies NAT and forwards the packets to the MN. No tunneling is needed for this stream since the path between the CN and the MN is segmented by nodes with NAT tables that can refer to the next hop address. This works fine for UDP traffic where there is no association of sender port number at the MN. However, in case of a TCP traffic, the MN also implements a NAT table eternally with entries for each session. This creates one more step of addresses translation, where IP addresses and port numbers of the received packets are replaced with those the MN had at the start time of the relevant traffic flows. Otherwise, the TCP session in a MN can break when receiving packets with a different IP address after a handover.

4.4.3 Update Mechanism

The proposed mechanism to handle the traffic continuity when handover and hence, changing the MN IP address is presented in this section.

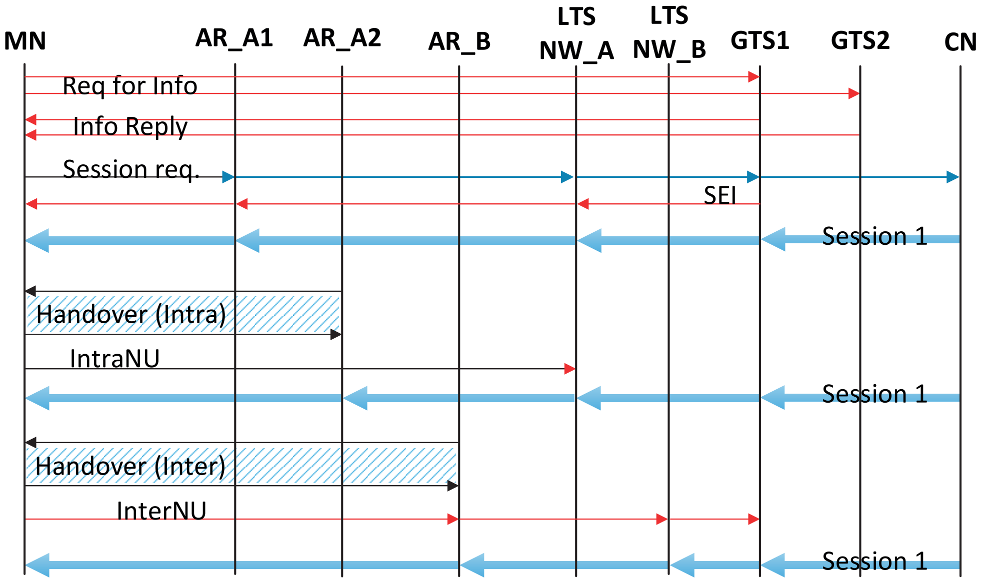

Figure 4.5: GTS selection and Handover using DiNAT. Red lines refers for DiNAT signaling messages, black for layer 2 handover and blue for the traffic stream.

4.4.3.1 The Process in the Mobile Node

Update messages are sent by the MN in case of a handover to populate the new address information to other relevant nodes. It is sent over UDP for the purpose of quick sharing of information. Two messages are employed depending on the type of the handover. For intra-mobility, where handover is between networks within the same administration domain, it is called Intra NAT Update (IntraNU), and Inter NAT Update (InterNU) for inter-mobility, where handover is between different domains. The recognition in between is achieved through the NID that we added to identify administration domains. The MN compares the NID supplied by the DHCP upon handover with that of the previously associated network and generates the appropriate update message accordingly.

The NAT update message carries information related to the previous and newly assigned MN address with the fixed session/index identifier useful for the update process in the GTS. It conveys also the MN MAC to the LTS to overcome the delay introduced by the ARP when trying to forward the first packet in the stream. The MN retrieves these information from its aforementioned repository that registers addresses and identifiers related to the running flows. It sends IntraNU only to the LTS and not further of the domain. InterNU are sent to the GTS, passing through the LTS. A MN might run sessions with multiple CNs, for which different GTSs might have been selected. In this case, a NAT update message is sent for each flow.

As mentioned earlier, these messages are sent by the MN over the new network after a successful link association and IP address assignment, during which ARP is preprocessed to precede packets forwarding as quick as possible.

4.4.3.2 The Process in the DiNAT Servers

An LTS receives both update types; Intra- and InterNU. It receives the first when a MN moves to a new network inside the same domain. In this respect, it uses the information carried in the update message to identify the concerned entry in its NAT table and modify it with the newly assigned address. After registering MN MAC and new IP address information, it starts to translate the exchanged packets with the updated IP address, and forwards the traffic to the MN in the new network.

An LTS receives the second update type (InterNU) when a MN moves to a new network but, in another administration domain. In this case, the LTS is the gateway of the new network. An InterNU is destined to the GTS, however, the LTS gets use of the carried information and forwards this message further to its destination. From this, the LTS creates a new NAT entry for the running session of the MN, since it is new in this network and its traffic will start to be received for the first time upon the update of the GTS and hence, re-route traffic to the new location.

A GTS receives the InterNU message to update its NAT table. Since this message passed also through the LTS, it is applied to NAT as well. A GTS looks up the combination of the session identifier besides the MN IP address and port numbers to find the correct entry. In this way, the GTS substitutes the registered LTS address with that of the new network. All next coming packets are translated to the new LTS address and sent forward.

4.4.4 DiNAT Servers Setup and Selection

We discuss here the location of the DiNAT servers (GTSs), the delivery of GTS nodes availability information to MNs, and the selection of a node among others as an anchor point for a stream.

A GTS node should be connected to the Internet, such that it does not lie under the administration of any service provider (local operators for example). For reason of load sharing and to validate the proposed approach, multiple nodes should be distributed through out the global network. In contrast to other solutions, a new GTS is selected at the start of an application, so a MN might utilize several nodes simultaneously other than only one.

The nature of this distribution is discussed in this dissertation, but only in a simplified manner. According to several trace route tests that we carried out from the networks of TU Ilmenau, a local DSL and a local UMTS provider to a variety of distributed global application servers (potential CNs), the RTT remains relatively small (<1ms to 9 ms) till the backbone routers in Germany (Further details on the architecture of the Internet can be found in [BDK04]). Accordingly, GTS nodes can be proposed to be distributed in sites of the backbone routers (8 in Germany). Such a setup can save extra signaling in the case of inter-domain mobility rather than when the GTS nodes would be installed far at the servers sites, where an RTT of the trace could reach more than 1000 ms.

To save MN signaling for looking up the current most suitable GTS nodes at the start of a new communication stream, we let the gateway of the network currently serves a MN to provide it with a list of the available GTS nodes through primitive request-reply signaling paradigm over UDP. Others means may also be useful, for example, to use DHCP, much like obtaining the default gateway and DNS addresses. This knowledge is assumed to be statically provided in the gateway since none of the nodes (neither the GTS nor the gateway) might have rapid dynamic changes in the connection topology.

As described earlier in the signaling procedure, a simple selection algorithm can choose one GTS among the provided list on bases of load and routing metrics. The later is calculated as the sum of the metrics of the GTS route to the MN from one side, and to the CN from the other side.

Any optimal route between any two nodes (the MN to the CN for example) usually is the one that has the lowest routing metrics, regardless of the used routing protocol. Similarly, less metrics between a MN, through a GTS, to the CN are preferred in the selection of the GTS, which represents a point in the middle that should be selected as close to the optimum route to the CN as possible. Another variant is a selection according to the minimum routing metrics of the GTS to CN path only. The last can be useful for highly moving nodes to save signaling higher GTSs due to frequent handovers when GTSs are to be deployed in multi-hierarchy as well.

4.5 Discussion

A main feature in the proposed handover system is flexibility. It has been designed as a framework, where the Controller module administrates the handover process and requests services from the Decider and the DiNAT-agent modules. The last two can be replaced with any other approaches that handle the same task namly; mathematical analysis of alternatives and address resolution (mobility management). Multi-criteria decisions and light weight signaling for quick traffic recovery are main features of the adopted modules as well.

Through the concept of DiNAT, this chapter emphasized mobility management as well. As have been seen from the previous chapter, mobility management solutions at the network layer adopt two main approaches. On one hand is the MIP concept, which must tunnel the traffic to wherever the MN is. On the other hand is the use of independent node addresses; for identification and for localization (routing), which requires to update a mapping sub-system between the two and all the communication peers, too.

DiNAT presents a new concept in the IP addressing. It segments the path hierarchically between the MN and the CN. The IP addresses in the higher segment facing the CN remains unchanged, while the lower segments down to the MN might be modified according to the mobility zone traversed by the MN. An update is signaled to the relevant DiNAT anchor point in the hierarchy upon a handover. Starting from the MN and up to the destined DiNAT anchor point, all other anchor points in the hierarchy make use of the update information to modify their mapping according to the MN's new location.

The use of hierarchical setup supports local and global mobility to reduce unnecessary signaling. Nevertheless, DiNAT emphasizes the role of NAT as dominant service in todays networks. To minimized the effect of the non-optimal routes problem and for purpose of load sharing, many DiNAT servers must be deployed in the Internet according to a reliable methodology and a simplified selection scheme as described in the previous section.

5 Simulations and Measurements

This chapter describes the simulation environment, the modeled networks and the scenarios used to evaluate the proposed framework. It provides details on the networks' topologies, their setup and the conducted scenarios in relation to the intended investigation. Measurements and evaluation are also provided, which we have published in [ARAS14, ARS15, ARS16, ARYS16, ARAS16b, ARAS16a] as well.

5.1 Simulation Environment

The simulation were successfully conducted using OMNeT++ v4.6 [Var] and the framework INET v2.3 [BMS+]. OMNeT++ is an extensible, modular, discrete-event, object-oriented C++ network simulation library and framework. It has a component architecture for simulation models. Models are assembled from components called modules, which can be connected to each other to form more complicated ones called compound modules. Itself, it is not a simulator, but rather provides the tools and the infrastructure to write simulations. INET is a framework and an open-source library for the OMNeT++ simulation environment. It benefits from the infrastructure, kernel and library provided by OMNeT++ to facilitate designing and validating protocols. It contains models for the Internet stack (IP, TCP , UDP, ect.), link layer protocols and mobility models.

5.2 Simulation Setup and Network Topologies

Three network topologies were simulated to evaluate the proposed approach. Some of the topologies were used in more than one scenario, as will be described in the next subsections.

Values related to the simulation environment are summarized in the tables 5.1, 5.2 and 5.5, while those related to the decision algorithm can be found in the tables 5.3 and 5.4. In each of the modeled topologies, networks have wireless AR and a gateway connecting to the Internet. The gateways provide also the DHCP service for each network, where different IP subnets are used (for ex., 192.168.1.0/24, 10.1.1.0/24 and 172.16.1.0/24).

5.2.1 Topology 1

As mentioned earlier, the decision of selecting a network among others and a time instant to execute handover (if required) is a critical factor in the overall performance of any mobility solution.

This topology provides a testbed to investigate the advantages of employing a multi-criteria decision algorithms and implementing a soft handover mechanism. For such measurements, we focus here on the events from the link layer's point of view. Handover issues at higher layers, like addressing, are investigated only with a limited scenario of the intra-mobility case, where a preliminary version of DiNAT were tested.

The topology involves three networks; Orange, Blue and Green, as shown in figure 5.1. They are different in terms of allocated radio channels, transmission power (and hence, coverage area), interference/communication range thresholds, bit rate and monetary service cost. The relative positions of the nodes are as shown in the figure. The coverage areas are different for each applied scenario therefore, their representation in the figure is general and does not express fixed values of power. Table 5.1 specifies the exact used values. Two mobility modules were employed (separately as required in the applied scenario) in this topology; a straight path (the weight dashed line) and a tractor path (the gray line). More on the traffic type and mobility model is provided in the scenarios section.

Each wireless interface of a MN is able to associate to only one network that lies in the same configured frequency/channel to ensure a switchover of connection to another wireless interface when a handover is decided. This basically applies to all the other topologies as well. The decision criteria and their weights are provided in tables 5.3 and 5.4.

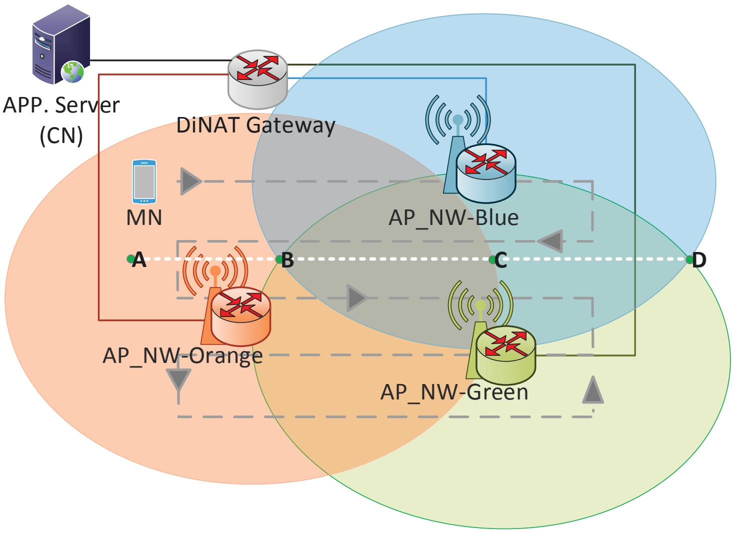

Figure 5.1: Topology 1: a testbed for MCDM algorithms and soft handover

5.2.2 Topology 2

IP address resolution upon a handover is a crucial factor that has direct impact on session continuity when a MN joins a new network/subnet. This topology focuses on handover from the addressing prospect. It enables the investigation of inter-mobility handover case, where a MN switches its connection between two networks/administration domains (see figure 5.2).

Furthermore, one of the networks is made supportive to our solution (network Blue), while the other (network Orange) is not (but only traditional NAT) in order to show the vulnerability of the proposed protocol and its compatibility with already-existing networks and protocols. All MNs are provided with two wireless interfaces. As described in Topology 1, each corresponds to only one of the networks.

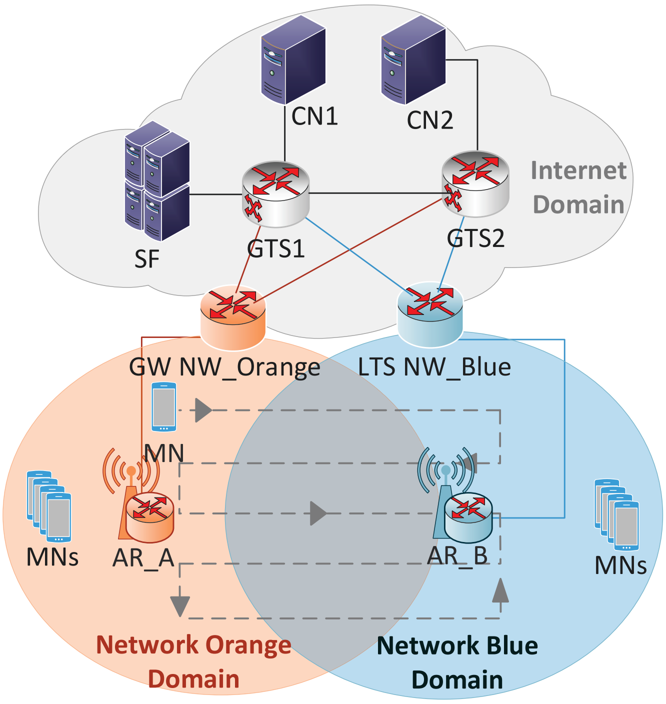

Figure 5.2: Topology 2: Inter-domain mobility

5.2.3 Topology 3

As described in the previous chapter, the proposed approach supports handover in intra- and inter-mobility environments, and with networks that do not support the DiNAT as well. The topology modeled in this section includes these possibilities. Similar to the previous topologies, the networks are different in its radio attributes and services. Areas of the same color represents the same administration domain. A network's coverage is illustrated as a circle.

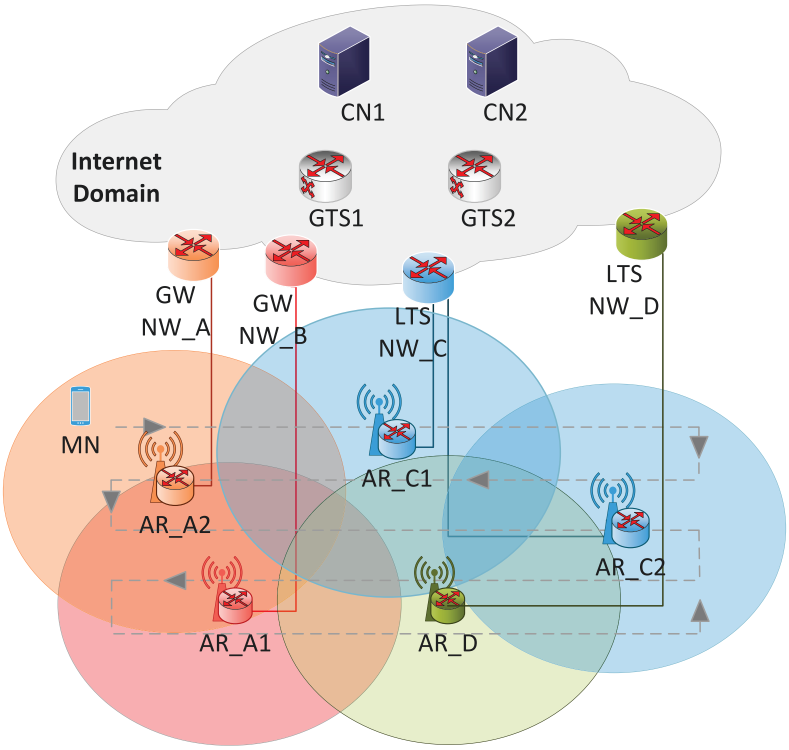

Figure 5.3: Topology 3: Inter- and Intra-mobility between 5 subnets in 3 domains. Each circle is a subnet in a domain represented by a the same color.

Scenarios of Topology 1	Parameters	Values
All scenarios	Access Technology	IEEE 802.11
	No. of wireless interf. of MN	3
	Channel of NW Orange	1
	Channel of NW Green	5
	Channel of NW Blue	3
	Propagation model	Free space
	Subnet of NW Orange	192.168.1.0/24
	Subnet of NW Green	172.16.1.0/24
	Subnet of NW Blue	10.1.1.0/24
	IP config.	DHCP
Scenario LDNiAT	Trans. power of NWs Orange	2 mW
	Trans. power of NWs Blue	2 mW
	Trans. power of NW Green	2.5 mW
	Bit rate of NW Orange	1 Mbps
	Bit rate of NW Green	2 Mbps
	Bit rate of NW Blue	11 Mbps
	Mobility model	Straight
	Mobility speed	1, 2, 5, 10 mps
	Video sending rate	2 Mbps
	Simulation time	240-24 s
Scenario MCDM	Trans. power of NW Orange	2.5 mW
	Trans. power of NW Green	2 mW
	Trans. power of NW Blue	1.8 mW
	Bit rate of NW Orange	1 Mbps
	Bit rate of NW Green	2 Mbps
	Bit rate of NW Blue	11 Mbps
	Mobility model	Tractor
	Mobility speed	5 mps
	Video sending rate	0.5, 1, 2 Mbps
	Simulation time	1000s
Scenario SMIP6	Trans. power of NW Orange	2mW
	Trans. power of NW Green	2mW
	Bit rate of NW Orange	2 Mbps
	Bit rate of NW Green	2 Mbps
	Mobility model	Tractor
	Mobility speed	1, 2, 4, 8, 10 mps
	Video sending rate	0.5, 2 Mbps
	VoIP Packet. Interval	20 ms
	Simulation time	2000-200 s

Table 5.1: Simulations parameters of the different scenarios of Topology 1

Scenarios of Topology 2	**Parameters**	**Values**
Scenario GDiNAT	Access Technology	IEEE 802.11
	Propagation model	Free space
	No. of wireless interf. of MN	2
	Channel of NW Orange	1
	Channel of NW Blue	3
	Subnet of NW Orange	192.168.1.0/24
	Subnet of NW Blue	10.1.1.0/24
	IP config.	DHCP
	Trans. power NW Blue	3 mW
	Bit rate NW Green	2 Mbps
	Bit rate NW Blue	12 Mbps
	Mobility model	Tractor
	Mobility speed	2 Mps
	Video sending rate	2 Mbps
	VoIP Packet. Interval	20 ms
	VoIP play out delay	5 ms
	Simulation time	900 s
	No. of seeds	3

Table 5.2: Simulations parameters of Topology 2

Relative criteria	**Importance**
RSS vs Delay	1
RSS vs Jitter	3
SNIR vs RSS	10
SNIR vs Cost	1
SNIR vs Delay	2
SNIR vs Jitter	6
Bit rate vs SNIR	5
Bit rate vs RSS	5
Bit rate vs Cost	5
Bit rate vs Delay	9
Bit rate vs Jitter	9
Cost vs RSS	10
Cost vs Delay	1
Cost vs Jitter	1
Jitter vs Delay	2

Table 5.3: Relative weighting of the criteria

Criteria	**Weights**
RSS	3
SNIR	6
Bit rate	9
Cost	1
Jitter	2

Table 5.4: Criteria weights

Scenarios of Topology 3	**Parameters**	**Values**
Scenario OpTest	Access Technology	IEEE 802.11
	Propagation model	Free space
	No. of wireless interf. of MN	5
	Channel of NW Blue C1	4
	Channel of NW Blue C2	5
	Channel of NW Orange	1
	Channel of NW Green	2
	Channel of NW Pink	3
	Subnet of NW Blue C1	10.1.0.0/16
	Subnet of NW Blue C2	10.2.0.0/16
	Subnet of NW Orange	192.168.1.0/24
	Subnet of NW Green	172.16.1.0/24
	Subnet of NW Pink	10.3.0.0/16
	IP config.	DHCP
	Trans. power of all NWs	2 mW
	Bit rate of all NWs	2 Mbps
	Mobility model	Tractor
	Mobility speed	5 Mps
	Video sending rate	0.5 Mbps
	VoIP Packet. Interval	20 ms
	VoIP play out delay	5 ms
	Simulation time	1000 s

Table 5.5: Simulations parameters of Topology 3

5.3 Scenarios and Measurements

Several simulation scenarios were conducted along this work. Each is relevant to one of the described topologies. However, more than one scenario can be applicable to the same topology, but each with a different setup. The scenarios are oriented in a sequence related to development phases of the work and the aforementioned publications. In order to keep a flow of information to the reader, measurements and evaluation are presented with each scenario as well. The scenarios are denoted with alphabetical annotations that should not be mixed with any mentioned abbreviation of protocols and mechanisms.

5.3.1 Scenario LDiNAT

5.3.1.1 Scenario LDiNAT: Description

This scenarios employs the aforementioned Topology 1. It investigates the flexibility of employing an MCDM decision algorithm and the advantages of applying soft handover. The MN moves in a constant speed with a linear mobility; similar to a straight road for example, from a predefined point within the coverage of network Orange, to which it is associated first, to a predefined one (along the road) in the overlap area between networks Blue and Green, where it then handovers to one of them. It moves back and forth on the road between the networks and through the overlap area of the three networks (point A and B in figure 5.1).

In this scenario, we employed a video stream traffic requested by the MN and provided by an application server. The MN requests the stream and starts moving from its initial position in network Orange. It moves in a constant speed over the defined pattern, while performing several handovers. We ran the simulation with different combinations of application sending rate and MN velocities, first with our solution (DiNAT) installed on the MN and second with MIPv6. In the second case, the MN is provided with a single wireless interface and a traditional decision strategy, which considers only the received signal strength as the decision criteria.

The criteria weights of the employed MCDM algorithm (AHP in this scenario) are configured here to extremely promote one of the networks (previous or next) to result in the earliest and the latest possible Handover Time Instant (HTI) in the simulated topology. Balanced weights were configured to result in intermediate different instants of handover time. Furthermore, it was interesting to observe how the two decision schemes (traditional and AHP), and due to the considered criteria, lead to handover into different networks. From addressing perspective, we compare also between the two cases by measuring the packet loss rate resulted by handover.

5.3.1.2 Scenario LDiNAT: Measurements and Evaluation

Through the simulations we ran in this scenario, we observed the effect of the decision algorithm on choosing the HTI and its impact on the throughput experienced by the MN at the application layer. The measurements shown here are selected samples out of a set of runs. An evaluation of the proposed address resolution scheme (despite the scenario is limited to intra-mobility) is provided in terms of the packet loss rate, which results from the latency introduced by the VHO while working to reroute an ongoing session to the new network and resume it running at the MN. With higher mobility speeds, we show also a scaled effect of the VHO on the packet loss rate.

In the MIPv6 case, a VHO is triggered only when the MN is on the edge of network Orange and is starting to loose its beacons (recognized as noise due to signal weakness). In the DiNAT case, a wide range of flexibility is introduced in shifting the HTI, from a minimum, which is as observed in the simulation as the point where the MN starts to have network Blue in range (point B in figure 5.1), to a maximum that lies at the edge (point C in the figure) of the connected network; network Orange. The HTI inside this range depends on the weights of the decision criteria (in the AHP), which we have set differently in each of the illustrated ten runs shown in figure 5.4.

In terms of throughput, figure 5.5 shows the effect of choosing a late/early HTI when the connected network (Orange) has poor QoS conditions. It shows a VHO with the earliest possible HTI (also shown in figure 5.4) when we set a high weight to the bit rate in the AHP, where network Blue was configured with a bit rate that is more than ten times that of network Orange. A traditional algorithm will wait till it starts to loose network coverage to initiate handover (the latest instant time in figure 5.4).

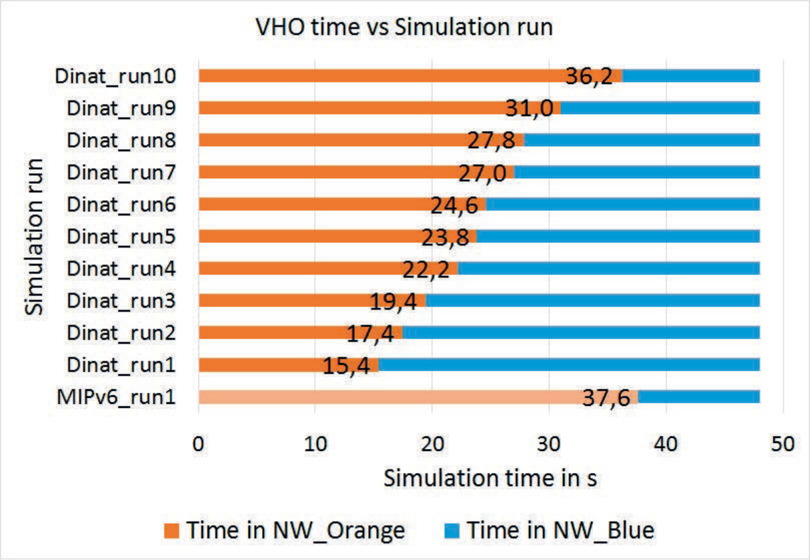

Figure 5.4: Handover time instants

Other than the HTI, figure 5.5 shows a handover of the MN to two different networks due to the use of different decision schemes in the tested cases (DiNAT and MIPv6). Case DiNAT shows a peak in throughput during handover due to short dual mode of connectivity through two interfaces, the previous and newly selected ones. In case MIPv6, the peak down is due to delay in recovering the traffic. The MN has to select between two networks when it decides a handover, either network Blue or Green, one with a less signal strength but a higher bit rate and the second is vise versa, respectively. When we set less importance to signal strength against bit rate, the MN selects network Blue using AHP and wins high throughput, while a traditional algorithm selects network Green with the higher power and suffers low rate, as can be seen in figure 5.5.

To investigate the performance of the address resolution phase, we set high conditions in term of QoS to all the networks in order to refer any dropped packets or delay only to the addressing issue. As shown in figure 5.6, the suggested solution shows a consistent performance, while MIPv6 shows an outage introduced due to the handover delay in this case that leads to lose packets till the traffic is recovered in the new network. A measured delay of about 2.65 seconds in average is introduced in MIPv6 case, while in the DiNAT case it is 0.01 seconds. The type of the decision algorithm applies no impact here since all overlapping networks are set to similar conditions for this test.

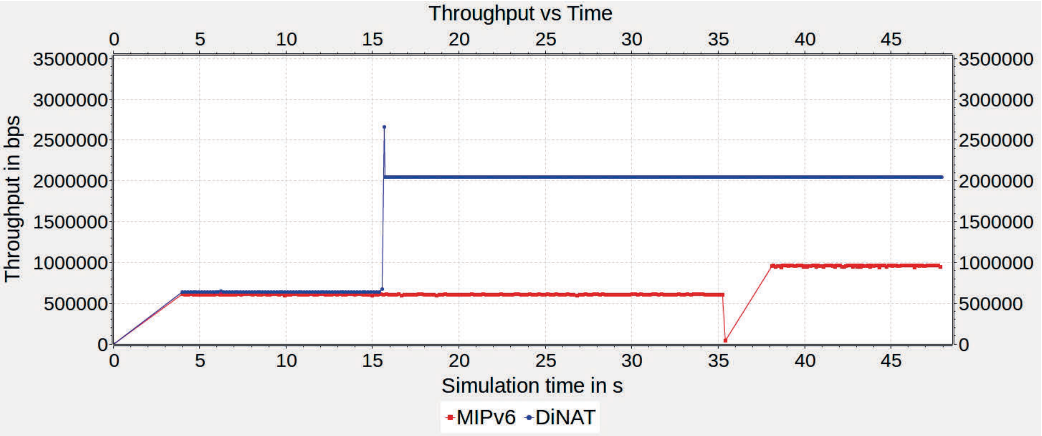

Figure 5.5: Throughput when handover to different networks at different time instants.

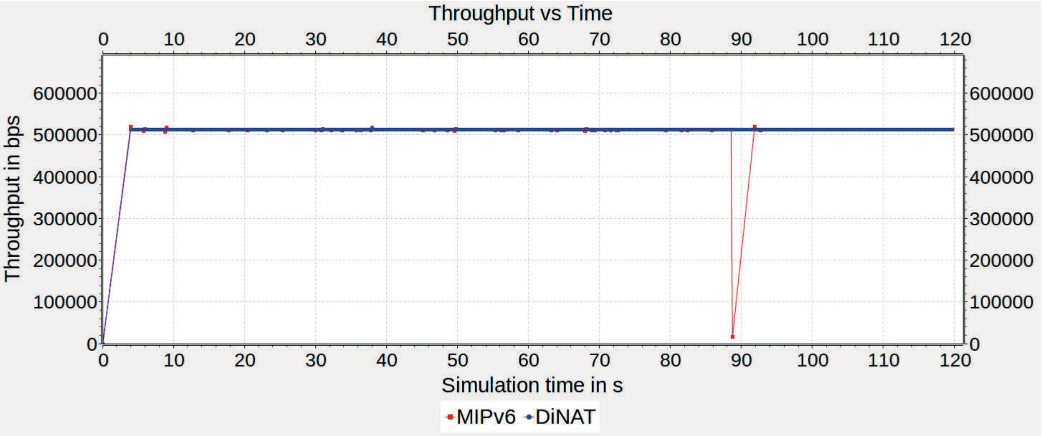

Figure 5.6: Throughput when handover to the same network at the same moment using different address resolution approaches

To present a scaled impact of the address resolution at handover, we repeat the runs with the different mobility speed values. Figure 5.7 illustrates a better performance for DiNAT. It shows, however, two DiNAT cases; Dinat v0 that implements hard handover, and the second employs soft handover. The last shows no lost packets along all the tested speeds.

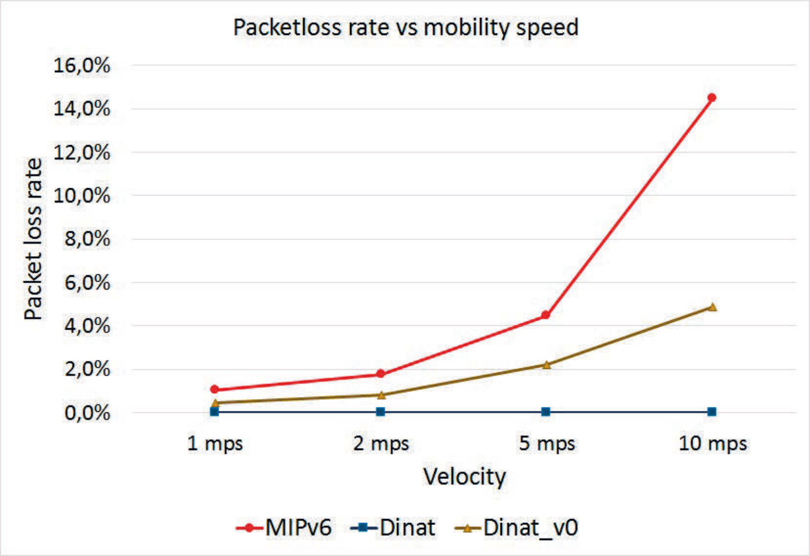

Figure 5.7: Packet loss rates

5.3.2 Scenario MCDM

5.3.2.1 Scenario MCDM: Description

As a significant phase in HO, decision making is examined in this scenario.The setup of Topology 1 is also used in this scenario to compare between a set of MCDM decision algorithms that are made available to be plugged in the Decider module. The MN moves on a tractor path with 2 mps constant velocity while receiving a video stream from the application server. The simulation is repeated for each of the considered algorithms, these are; WSM, WPM, VIKOR, and GRA. The performance is evaluated in terms of the average decision criteria values as experienced by the MN during each run. The criteria weights were configured to dominate the network's bandwidth.

5.3.2.2 Scenario MCDM: Measurements and Evaluation

We ran the simulated scenario several times under the same conditions while employing a different decision algorithm each time. An additional algorithm we call Trad is added as well, which is a traditional decision based on RSS only. The average criteria values were calculated for the MN at the end of each simulation, in which it remained connected to certain networks for a specific amount of time based on the decisions taken by the employed HO decision algorithm.

The evaluation is carried out on basis of the decision criteria considered in the algorithm, the amount of the received bytes at the application layer and how often a HO was performed, see figure 5.8. Some of these are cost criteria; based on the concept of the lower, the better, and some are gain ones; on basis of the higher, the better.

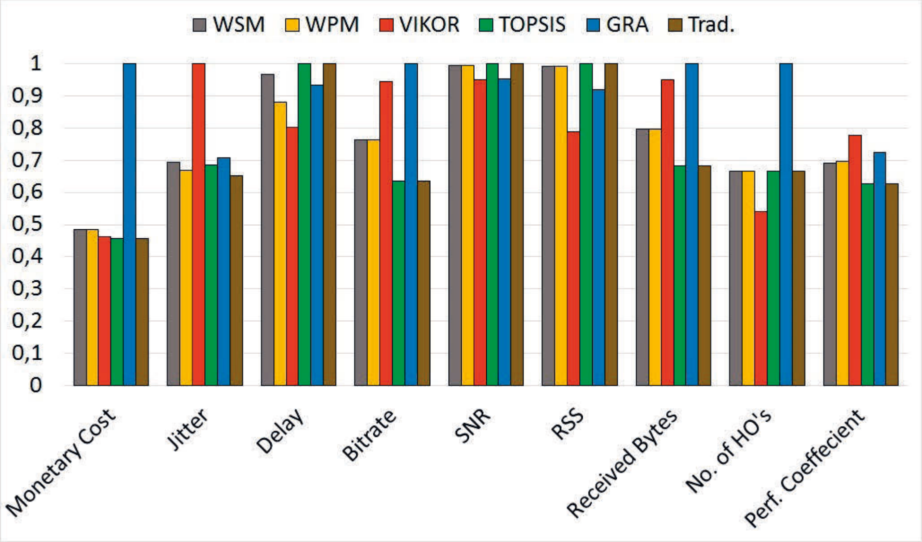

Figure 5.8: The normalized average values of the evaluation criteria as experienced by the MN for the considered decision algorithms

A performance coefficient is calculated for each algorithm as the sum of products of the decision criteria measured values and its weights, plus the received bytes and the number of HOs (all normalized values), However, cost- and gain-based criteria were taken under consideration for this calculation. As the figure shows, GRA and VIKOR present an advantage in terms of experienced bit rate and hence, the number of received bytes, however, with relatively lower RSS and SNIR. All other algorithms show approximately similar behavior. These results for GRA and VIKOR correlate with the outcome in [SNW06] and [LBS15] respectively. This scenario shows the flexibility gained when utilizing different MCDM-based algorithms in HO decisions. Therefore, the Decider module is implemented as a framework with the ability to integrate (but not limited to) any of these tested algorithms.

5.3.3 Scenario SMIP6

5.3.3.1 Scenario SMIP6: Description

This scenario considers testing MIPv6 when soft handover is added. Topology 1 suits for this scenario but, substituting the nodes AP_NW_Orange and AP_NW_Green with home agent and foreign agent nodes respectively. Network Blue is unnecessary to be included in this scenario so it has been omitted from the topology for this scenario. The MN moves on a tractor path as illustrated in the simulation topology. A traditional decision algorithm is employed for both case; hard and soft handover. The simulation was repeated for each of these two cases. Soft handover is enabled through the employment of our link layer controller module; the Controller. In each case, video streaming and VoIP traffic were tested separately.

5.3.3.2 Scenario SMIP6: Measurements and Evaluation

Hard and soft handovers are compared to each other using QoS-related parameters, which were collected while running each of the used applications. Figures 5.9 and 5.10 illustrate the packet loss rate for the video and VoIP traffic respectively. We notice a significant enhancement presented by soft handover when enabled for MIPv6, especially at higher velocities. However, for the VoIP application, the packets are not sent as a stream, but rather as talk spurs therefore, lately arrived packets might also be considered as lost in this case.

The Mean Openion Score (MOS) measured by OMNeT++ for the VoIP application shows relatively similar performance at low speeds for both HO schemes, but better values can be achieved at higher speeds by the Soft HO, as can been noticed in figure 5.11. Handover in Soft HO case takes place while being in the coverage overlap area between the two networks and the switching is executed only when the second connection has been really created (make-before-break). In Hard HO case, handover is performed only when the MN is on the edge of the connected network, where it starts associating its single wireless interface to the upcoming network after notifying that no more beacons are recognized, but as noise.

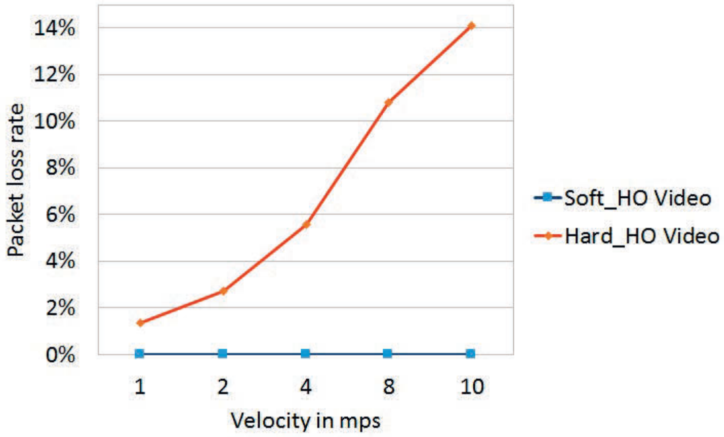

Figure 5.9: Packet loss rate in the Video traffic case

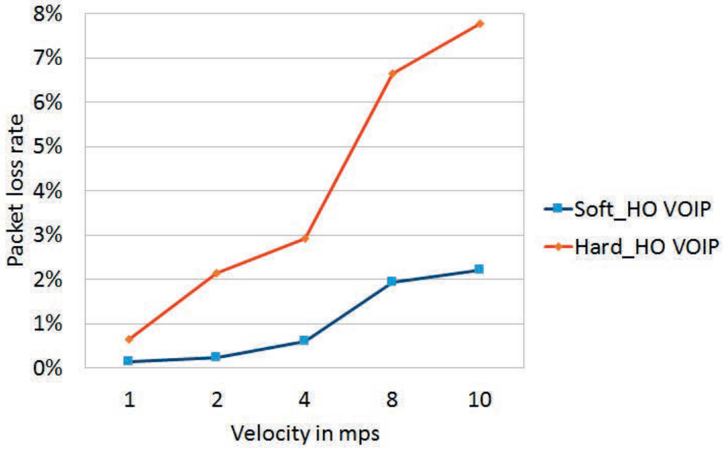

Figure 5.10: Packet loss rate in the VoIP traffic case

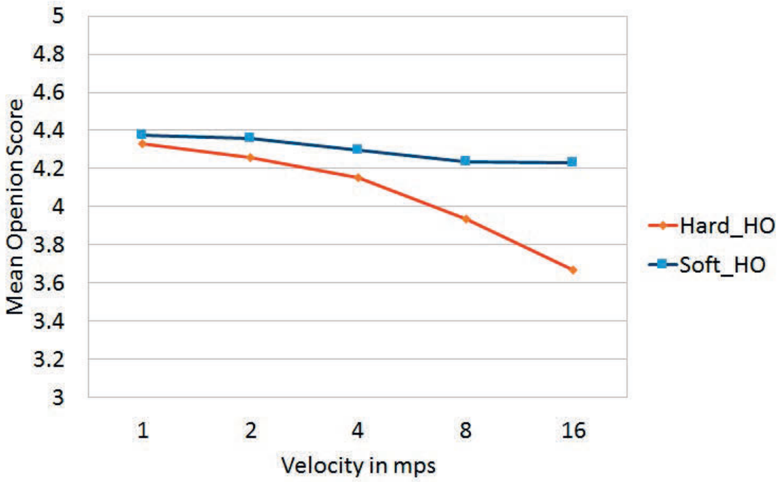

Figure 5.11: MOS of the VoIP traffic case

5.3.4 Scenario HDiNAT

5.3.4.1 Scenario HDiNAT: Description

In this scenario, Topology 2 was used to evaluate the proposed DiNAT approach in an inter-mobility scenario. The network model assumption is relaxed in this case, which is more close to real setup of networks; service providers/operators might have gateways to the Internet independent of each other. This scenario aims at the evaluation of different handover approaches from the addressing perspective with soft switching of traffic over multi-interfaces, however, hard HO case was also tested as a reference of basic behavior.

Three cases were tested in this scenario; case one employing our suggested solution, case two employs MIPv6, but with hard handover and case three is MIPv6 with soft handover enabled. We must notify here that MIPv6 model provided in OMNeT++/INET framework supports only hard handover. Therefore, we enable soft handover by implementing our management module described in 4.1. Details on the conducted modification were published in [ARAS16a]. A MN starts a real-time application with two servers at the same time and moves around in the meanwhile. The scenario is repeated for each of the tested cases; first using Video stream traffic and again using VoIP for each. The routing metrics of the gateway, LTS and GTS's are set (statically), such that a different GTS node shall be selected as an anchor point for each session in the DiNAT case.

5.3.4.2 Scenario HDiNAT: Measurements and Evaluation

Different GTS nodes were used for the two traffics a MN has ran. The GTS selection algorithm considers the routing metrics, which were configured statically to lead to this selection, and hence emphasizing the concept of DiNAT (details on GTS selection scheme are mentioned in section 4.4.4). In MIPv6, the anchor point is always the home agent. We ran the simulations with a MN's constant mobility speed and observed handover latency, packet loss rate and end-to-end delay introduced to the running traffic due handover events. The handover latency is measured at the MN and it is defined as the time difference between the last packet received through the old IP address, and the first one received with the new address after handover.

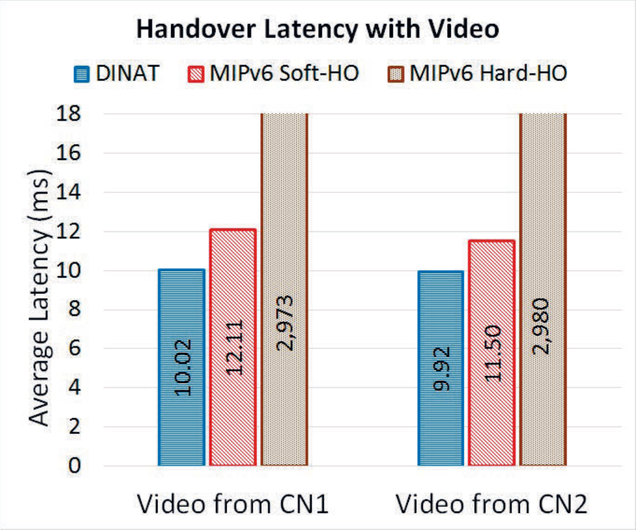

Figure 5.12: Handover latency for the video traffic

Figure 5.12 shows the measurements related to the video traffic. We cannot rely on VoIP for this parameter since it is not a continuous stream in its nature. DiNAT shows better results because, unlike MIPv6, packets do not have to be forwarded to a HA. Handover latency is less in DiNAT case for the reason of early learning of the MAC addresses also, because MACs of the MN and the network gateway are exchanged during link association phase.

In terms of packets loss, both solution that use soft handover show no losses when running the video stream, while the hard HO version of MIPv6 shows a 2.8 percentage for the video traffic case. This difference can be referred to the basic concept of soft versus hard switching.

Signaling and update messages to reach anchor points are also very effective parameters in reducing the amount of lost packets if they can be achieved faster than the sending rate of the running traffic. Packet loss rate in VoIP is more sensitive to delay, where even when a packet has been successfully received, it might be irrelevant for an ongoing conversation if it arrives too late. Such packets are considered as lost (see figure 5.13). Tunnel establishment adds a considerable delay and hence, results in discarding lately arrived packets. In DiNAT case, a significant improvement was observed in terms to this parameter.

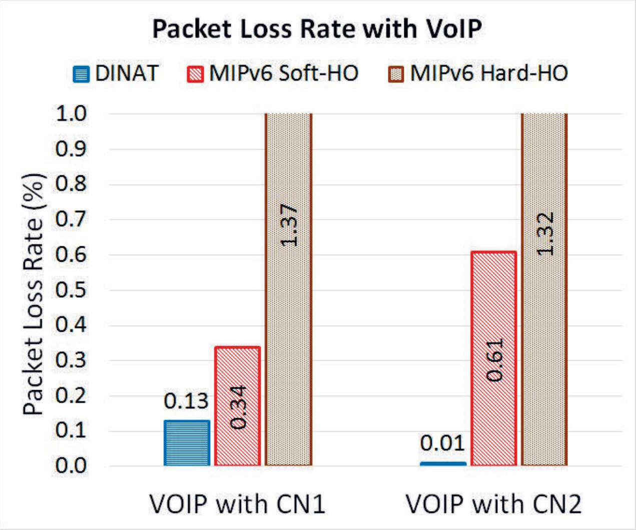

Figure 5.13: Packet loss rate for VoIP traffic

Mean opinion score (MOS) for VoIP was observed, however, at relatively low speed (2 mps) used for mobility, the MOS was almost the same for the examined approaches (see figure 5.14). This has been seen in a previous scenario and topology to have advantage when implementing DiNAT at higher velocities.

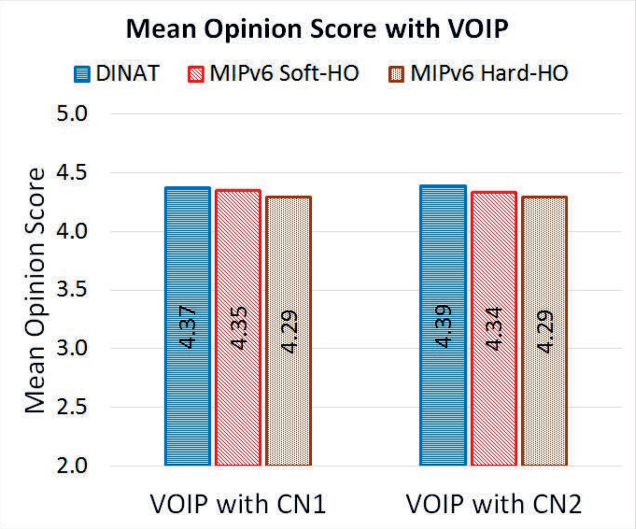

Figure 5.14: MOS of the VoIP traffic

We measure also the end-to-end delay for the whole transmission duration of the traffic. Such a parameter is not so relevant to handover performance, but rather to the path between the MN and the CN passing through the anchor point. Synchronization between sender and receiver is not considered since end-to-end delay is only a measurement in the simulation for performance evaluation and has not been used as input to any further process. Figure 5.15 shows this delay for both considered traffic applications.

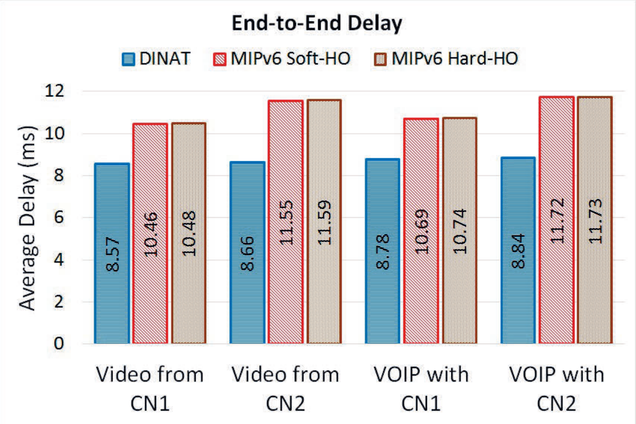

Figure 5.15: End-to-End delay for both traffics

5.3.5 Scenario OpTest

Topology 3 (figure 5.3) was used to run a scenario that can apply handover cases of inter- and intra-mobility in a combination of different networks, where some do support DiNAT, while others have no knowledge of its implementation. The goal of this scenario is not to compare any measurements since the former scenarios demonstrate the main goals, and it would be hard to argue results precisely when influenced by the implementation of many modules. However, the goal is rather intended here to test and ensure an error-free behavior of DiNAT when applying the mentioned cases together.

The MN moves on a tractor path and run real-time applications to traverse the coverage areas and switch between the networks. Two streams of video traffic run between the MN and two application servers in the global network denoted as CN1 and CN2. The simulation is repeated using two VoIP sessions similarly, each between one of the CNs and the MN. The simulations ran successfully across the simulation time. Handover was performed softly 18 times in the overlap areas during each of the runs. The CNs were able to send approximately 6.3 MB of video for each, which were all received by the MN with no lost packets. An MOS of 4.4 was experienced in both of the VoIP applications.

5.4 Discussion

According to the presented measurements, we can briefly refer the achieved enhancements to the soft switching, the multi-criteria decisions and the simple address resolution scheme.

The first creates and configures a new interface/path due deciding to the selected network before actually switching any running traffic from a previous interface. It shows enhancements in both tested handover approaches; MIPv6 and the proposed DiNAT.

In multi-criteria decisions, a more sophisticated handover is enabled, where a variety of criteria can contribute in decision making and a user of a MN can apply certain preferences through selecting the necessary configuration set (the service profile).

The simplistic signaling scheme to update entries of certain server/s that is providing NAT has fasten the recovery of the stream upon a handover. The employment of DiNAT in two layers save unnecessary signaling in intra-mobility case, while a global DiNAT can handle mobility between different administration domains. Another factor that facilitates faster resume of streams after a handover is the early learning of MAC addresses. A MN learns the gateway's MAC upon association, while gateways learn that of a MN upon the receipt of NAT update message. It was in our interest to test scenarios of real-time application that are sensitive to delay, in order to evaluate the proposed approach including the described features.

6 Conclusions and Outlook

This chapter concludes the work that has been presented throughout this dissertation. An outlook towards open issues and future possible extensions of this work is also given here.

6.1 Conclusions

This dissertation highlighted the handover between heterogeneous networks, which is expected to be a frequent process in today's communication and usage scenarios. The management of the complete process has been addressed, with considerations to quality of service parameters. QoS has been supported through performing soft switching between the device's interfaces, and adopting a light weight signaling scheme for address resolution. The handover decision algorithm engages users preferences as well rather than only traditional power-related criteria. The identity of a MN is maintained through the employment of the DiNAT solution.

DiNAT applies no modification on the communication parties, but rather uses DiNAT-enabled nodes as anchor points to hide the mobility problem from the other end. Global and local mobility are managed through a hierarchical employment of global and local DiNAT-nodes. Using the NAT concept, the traffic path is split into segments, where each is identified by a different IP address, which represents the MN side for that traffic. The path segment affected by the mobility zone (local/global) is prone to the update mechanism inside the related DiNAT server. For purpose of load sharing and in order to avoid the non-optimal route issue that exists in MIP-based solutions, many DiNAT-enabled nodes should be deployed globally, where no single operator can have exclusive administration rights to it. A selection mechanism chooses the one that is closest to the optimal path between the MN and the CN on basis of routing metrics and load. Considering the addressing space, DiNAT supports the used private addresses and leverages NAT to include mobility management service as well.

An overview on handover phases have been provided in chapter 2 with focus on decision strategies and algorithms. The main requirements to achieve our desired type of handover were also addressed in this chapter. Accordingly, the design of a handover model has been given with details on three sub-modules responsible of achieving the specified tasks. These namely are the Controller as a handover strategy manager and an overlay that coordinates the roles of the interfaces in the access layer, the Decider as a decision framework, where the desired mathematical algorithm can be plugged in to deliver MCDM-based decisions, and the DiNAT-agent that implements the introduced DiNAT approach and maintains correct interface's information in the device's routing table.

Following an overview on addressing generally in IP-based networks, more insight on the handover from a node's addressing perspective has been given in chapter 3. Different approaches with respect to the TCP/IP layers have been given with highlight on the most recognized ones. However, layer 3-based solutions were deeper investigated since a desired goal in this work is to apply minimum modifications to existing networks and protocols. To conduct a faster handover, cross-layer information regarding measurements from the access layer were necessary to be involved in the solution in favor to the handover strategy.

Further details on the proposed handover system and its modules have been specified in chapter 4. Here, the handover strategy has been described, which is managed by the Controller to carry out a soft and stable handover process. A set of MCDM algorithms is made available for implementation by the Decider module. However, complex weighting of criteria might be a non-reliable and an undesirable task to be performed by the user of a smart device, therefore, several service profiles are made available for selection to implement the usage policy desired by the user through applying the corresponding algorithm and weights. The DiNAT-agent implements the suggested DiNAT mobility management solution inside the MN. However, other nodes in the network take part also in the DiNAT process.

In the simulations and measurements part in chapter 5, three main topologies applicable to five scenarios have been described. The goals of conducting these scenarios are to show the performance with respect to the flexibility added by MCDM-based decisions in controlling handover execution time instants, to observe the optimization of QoS parameters due to choosing a different set of networks along the simulation runs when using each of the configured MCDM algorithms, and to evaluate the packet loss rate and handover latency when applying different combinations of hard-HO with soft-HO using DiNAT and MIPv6. Real-time applications of video stream and VoIP traffic have been used in the tested scenarios.

The proposed handover management system has achieved soft handover, low latencies and packet loss rate during handover. This is due to the support of local/global mobility, implementing soft switching between interfaces, and the employment of an anchor point selected close to the optimal path to the CN. The system uses layer 2 events and measurements for decisions and improves traffic recovery upon new associations. The system meets today's usage scenarios by considering user preferences in multi-criteria decisions, MN speed, data rate throttling and timers for stable handover. It represents a quick solution for today's IPv4-based networks and cooperates with NAT to prolong the life time of its address space. However, it is also applicable to IPv6-based network, where the main idea of using NAT is to segment the traffic path regardless of the protocol version. For realization in actual networks, it requires the deployment of DiNAT-enabled proxies in the Internet to support global mobility, nevertheless local mobility can be supported as well to save signaling by implementing DiNAT on gateways of network operators.

In addition to the aimed UDP-based applications, session continuity of TCP is enabled as well through a local translation of the MN IP address with the one used upon a transport association. Any new acquired IP address will remain invisible to the application and it is up to the DiNAT-agent to maintain the mapping.

The solution is also applicable to wireless stationary nodes, which may wish to switch connectivity in cases of degradation in network conditions for example. Decisions of horizontal handovers can be easily also included in the process, where each interface can forward the list of the discovered access points to the Controller through a request-to-connect inquiry, too.

6.2 Outlook

Along the progress of this work to investigate the specified problems and their corresponding solutions, several issues have been addressed for further investigations to enhance the proposed approach and make it better suited for realization.

A part of the assumptions is to have the wireless interfaces available on the device active simultaneously to maintain a sustained monitoring of the environment and the conditions of available networks. Despite the witnessed improvements in the field of batteries manufacturing to support high energy consumption, especially with the trend of today's big-display smart phones, better energy management policies might save amounts of energy. For instance, to switch interfaces ON and OFF with the initiation and during the running of specific applications. This function can be extended inside the Controller, which has a central access to all handover-related functions inside the interfaces.

Concerning the decision algorithms, more investigations might be useful to define a best suited MCDM algorithm for each traffic class. The desired algorithm can be easily configured to be employed as soon the application defined constrains are met. The relevant criteria could be different in each case as well. Normalization functions used inside the MCDM algorithms might be implemented to suite specific usage polices, like to select the lowest sufficient parameters rather than going always for the highest value.

The tested scenarios dealt with a typical setup of application servers in the global network, while the mobile node lies behind a NAT/firewall in a private network. For scenarios where an application server is connected in the private network and might be mobile as well, two problems are associated; the reachability to the server from outside and the dynamic adaptation to changes in address due to mobility. Methods of port forwarding, where specific transport destination ports are associated to specific hosts inside the local network, can be employed and further developed to use the same described update scheme in DiNAT.

The suggested solution considered no security features, but rather handled the issue from a pure layer-3 perspective. Additional efforts on security are necessary if the solution wanted to be realized. Other mechanisms to guarantee QoS rather than only supporting it are also useful to achieve better user's satisfaction.

Bibliography

[ABG14] A. Ahmed, L. M. Boulahia, and D. Gaiti. Enabling Vertical Handover Decisions in Heterogeneous Wireless Networks: A State-of-the-Art and A Classification. *IEEE Communications Surveys Tutorials*, 16(2):776–811, Second 2014.

[Aln12] E. Alnasouri. *QoS-aware Mobility Management in IP-based Communication Networks*. PhD thesis, Technische Universitaet Ilmenau, 2012.

[ARAS14] A. Al-Rubaye, A. Aguirre, and J. Seitz. Poster: Address Resolution for Vehicular Communications in Heterogeneous Environments. In *IEEE Vehicular Networking Conference (VNC)*, pages 123–124, Dec 2014.

[ARAS16a] A. Al-Rubaye, A. Aguirre, and J. Seitz. Enabling Soft Vertical Handover for MIPv6 in OMNeT++. In *Proceedings of the OMNeT++ Community Summit*, 2016.

[ARAS16b] A. Al-Rubaye, A. Aguirre, and J. Seitz. Hierarchical Dynamic Index NAT as a Mobility Solution. In *The 3rd IEEE IDAACS Symposium on Wireless Systems within the IEEE International Conferences on Intelligent Data Aquisition and Advanced Computing Systems: Technology and Applications*, 2016.

[ARS15] A. Al-Rubaye and J. Seitz. Dynamic Index NAT as a Mobility Solution in OMNeT++. In *Proceedings of the OMNeT++ Community Summit*, 2015.

[ARS16] A. Al-Rubaye and J. Seitz. A Cross-Layer Mobility Management with

Multi-Criteria Decision Making. In *Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 821–826, July 2016.

[ARYS16] A. Al-Rubaye, Y. Yeryomin, and J. Seitz. Evaluation of MCDM-based Handover Decision Algorithms in OMNeT++. In *The 3rd IEEE IDAACS Symposium on Wireless Systems within the IEEE International Conferences on Intelligent Data Aquisition and Advanced Computing Systems: Technology and Applications*, 2016.

[BDK04] A. Budke and A. Pott D. Kanwischer. *Internetgeographien: Beobachtungen zum Verhältnis von Internet, Raum und Gesellschaft (Erdkundliches Wissen).* Franz Steiner Verlag, 2004.

[BHSM05] M. Buddhikot, A. Hari, K. Singh, and S. Miller. MobileNAT: A New Technique for Mobility Across Heterogeneous Address Spaces. *Mobile Networks and Applications*, 10(3):289 – 302, 2005.

[BMS$^+$] Z. Bojthe, L. Meszaros, B. Seregi, R. H., and A. Varga. INET Framework.

[BMvB11] M. Bagnulo, P. Matthews, and I. van Beijnum. Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers. Request for Comments 6146, Internet Engineering Task Force, April 2011.

[CH92] Sh. Chen and Ch. Hwang. *Fuzzy Multiple Attribute Decision Making.* Springer-Verlag, 1992.

[CM04] A. Calvagna and G. Di Modica. A User-Centric Analysis of Vertical Handovers. In *The Second ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots*, pages 137–146, 2004.

[Dia10] A. Diab. *Mobility Management in IP-Based Networks. Analysis, Design, Programming and Computer-Based Learning Modules.* PhD thesis, Technische Universität Ilmenau, 2010.

[DS14] A. Dutta and H. Schulzrinne. *Mobility Protocols and Handover Optimization: Design, Evaluation and Application.* John Wiley & Sons, 2014.

[Edd04] W. M. Eddy. At What Layer Does Mobility Belong? *IEEE Communications Magazine*, 42(10):155–159, Oct 2004.

[ES08] F. Evers and J. Seitz. REACH: A Roaming-Enabled Architecture for Multi-Layer Capturing. In *IEEE Wireless Communications and Networking Conference WCNC*, 2008.

[FFLM14] D. Farinacci, V. Fuller, D. Lewis, and D. Meyer. LISP Mobile Node. Internet-Draft, 2014.

[FFML13] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis. The Locator/ID Separation Protocol (LISP). Request for Comments 6830, Internet Engineering Task Force, January 2013.

[FGM$^+$99] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. Hypertext Transfer Protocol – HTTP/1.1. Request for Comments 2616, Internet Engineering Task Force, June 1999.

[Fis67] P.C. Fishburn. Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments. *Operations Research*, 15:537–542, 1967.

[FK12] S. Fernandes and A. Karmouch. Vertical Mobility Management Architectures in Wireless Networks: A Comprehensive Survey and Future Directions. *IEEE Communications Surveys Tutorials*, 14(1):45–63, First 2012.

[For03] B. A. Forouzan. *TCP/IP Protocol Suite*. McGraw-Hill, 2003.

[FRHB13] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure. TCP Extensions for Multipath Operation with Multiple Addresses. Request for Comments 6824, Internet Engineering Task Force, January 2013.

[GK11] M. K. Gondara and S. Kadam. Requirements of Vertical Handoff Mechanism In 4G Wireless. *International Journal of Wireless and Mobile Networks (IJWMN)*, Volume 3(No. 2):18 – 27, April 2011.

[GLD+08] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil. Proxy Mobile IPv6. Request for Comments 5213, Internet Engineering Task Force, August 2008.

[GSRM05] L. Galingo-Sanchez and P. Ruiz-Martinez. QoS and Micromobility Coupling: Improving Performance in Integrated Scenarios. *European Journal for the Informatics Professional*, 5:49 – 55, 2005.

[HI10] A. Huszak and S. Imre. Eliminating Rank Reversal Phenomenon in GRA-Based Network Selection Method. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–6, May 2010.

[HY81] C.L. Hwang and K. Yoon. *Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag, 1 edition, 1981.

[IS15] J. Iyengar and I. Swett. QUIC: A UDP-Based Secure and Reliable Transport for HTTP/2. Internet-Draft, 2015.

[JCV+13] J. Jailton, T. Carvalho, W. Valente, C. Natalino, R. Frances, and K. Dias. A Quality of Experience Handover Architecture for Heterogeneous Mobile Wireless Multimedia Networks. *IEEE Communications Magazine*, 51(6):152–159, June 2013.

[JPA04] D. Johnson, C. Perkins, and J. Arkko. Mobility Support in IPv6. Request for Comments 3775, Internet Engineering Task Force, June 2004.

[KKP08] M. Kassar, B Kervella, and G. Pujolle. An Overview of Vertical Handover Decision Strategies in Heterogeneous Wireless Networks. *Computer Communications*, 31(10):2607 – 2620, 2008.

[KM15] S. Kumbalavati and J. Mallapur. A Survey on Vertical Handover in Heterogeneous Network. *International Journal of Information and Futuristic Research*, 2:3028–3033, 2015.

[Koo05] R. Koodli. Fast Handovers for Mobile IPv6. Request for Comments 4068, Internet Engineering Task Force, July 2005. Obsoleted by RFC 5268.

[KPKP11] Y. Kima, S. Packa, Ch. Kanga, and S. Parkb. An Enhanced Information Server for Seamless Vertical Handover in IEEE 802.21 MIH Networks. *Computer Networks*, 55:147–158, 2011.

[KTI03] M. Kunishi, F. Teraoka, and M. Ishiyama. LIN6: A Solution to Mobility and Multi-Homing in IPv6. Internet draft, 2003.

[Lam07] T. Lammle. *CCNA : Cisco Certified Network Associate Study Guide*. John Wiley & Sons, September 2007.

[LBS15] M. Lahby, S. Baghla, and A. Sekkaki. Survey and Comparison of MADM Methods for Network Selection Access in Heterogeneous Networks. In *7th International Conference on New Technologies, Mobility and Security (NTMS)*, pages 1–6, 2015.

[LHM10] H. Lee, Y. Han, and S. Minl. Network Mobility Support Scheme on PMIPv6 Networks. *International Journal of Computer Networks and Communications*, 2:206 – 213, 2010.

[LL11] F. Y. Leu and K. C. Liang. A Location-Based Handoff Scheme Based on IEEE 802.21 in Heterogeneous Wireless Networks. In *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2011 Fifth International Conference on*, pages 571–576, June 2011.

[LSK+09] S. Lee, K. Sriram, K. Kim, Y. H. Kim, and N. Golmie. Vertical Handoff Decision Algorithms for Providing Optimized Performance in Heterogeneous Wireless Networks. *IEEE Transactions on Vehicular Technology*, 58(2):865–881, Feb 2009.

[MB98] D. A. Maltz and P. Bhagwat. MSOCKS: an Architecture for Transport Layer Mobility. 3:1037–1045 vol.3, Mar 1998.

[MIH09] IEEE Standard for Local and metropolitan area networks - Media Independent Handover Services. *IEEE Std 802.21-2008*, Jan 2009.

[MNJH08] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson. Host Identity Protocol. Request for Comments 5201, Internet Engineering Task Force, April 2008.

[MYO03] H. Matsuoka, M. Yoshimura, and T. Ohya. A Robust Method for Soft Ip Handover. *IEEE Internet Computing*, 7(2):18–24, Mar 2003.

[NGH10] P. Nikander, A. Gurtov, and T. R. Henderson. Host Identity Protocol (HIP): Connectivity, Mobility, Multi-Homing, Security, and Privacy over IPv4 and IPv6 Networks. *IEEE Communications Surveys Tutorials*, 12(2):186–204, Second 2010.

[NHH06] N. Nasser, A. Hasswa, and H. Hassanein. Handoffs in Fourth Generation Heterogeneous Networks. *IEEE Communications Magazine*, 44(10):96–103, 2006.

[NJP+13] A. R. Natal, L. Jakab, M. Portolés, V. Ermagan, P. Natarajan, F. Maino, D. Meyer, and A. C. Aparicio. LISP-MN: Mobile Networking Through LISP. *Wireless Personal Communications*, 70(1):253–266, 2013.

[NSS+11] P. Nevesa, J. Soaresb, S. Sargentob, H. Piresb, and F. Fontes. Context-Aware Media Independent Information Server for Optimized Seamless Handover Procedures. *Computer Networks*, 55:1498–1519, 2011.

[OBS+08] A. De La Oliva, A. Banchs, I. Soto, T. Melia, and A. Vidal. An Overview of IEEE 802.21: Media-Independent Handover Services. *IEEE Wireless Communications*, 15(4):96–103, Aug 2008.

[OT04] S. Opricovic and G. Tzeng. Compromise Solution by MCDM Methods: A Comparative Analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, 156(2):445 – 455, 2004.

[OT07] S. Opricovic and G. Tzeng. Extended VIKOR Method in Comparison with Outranking Methods. *European Journal of Operational Research*, 178:514–529, 2007.

[PCT03] C. Politis, K. A. Chew, and R. Tafazolli. Multilayer Mobility Management For All-ip Networks: Pure Sip Vs. Hybrid Sip/mobile Ip. In *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE*

Semiannual, volume 4, pages 2500–2504 vol.4, April 2003.

[Per02] C. Perkins. IP Mobility Support for IPv4. Request for Comments 3344, Internet Engineering Task Force, August 2002.

[Pol96] G. P. Pollini. Trends in Handover Design. *IEEE Communications Magazine*, 34(3):82–90, Mar 1996.

[RAD13] N. Rajule, B. Ambudkar, and A. Dhande. Survey of vertical Handover Decision Algorithms. *International Journal of Innovations in Engineering and Technology (IJIET)*, 2:362 – 368, 2013.

[RPB+12] C. Raiciu, C. Paasch, S. Barre, A. Ford, M. Honda, F. Duchene, O. Bonaventure, and M. Handley. How Hard Can It Be? Designing and Implementing a Deployable Multipath TCP. In *9th USENIX Symposium on Networked Systems Design and Implementation*, 2012.

[RSC+02] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. SIP: Session Initiation Protocol. Request for Comments 3261, Internet Engineering Task Force, June 2002.

[Saa08] T. L. Saaty. Decision Making with the Analytic Hierarchy Process. *Int. J. Services Sciences*, 1(1):83 – 98, 2008.

[SCEB08] H. Soliman, C. Castelluccia, K. ElMalki, and L. Bellier. Hierarchical Mobile IPv6 (HMIPv6) Mobility Management. Request for Comments 5380, Internet Engineering Task Force, October 2008.

[Sch03] J. Schiller. *Mobilkommunikation.* Pearson Studium, 2003.

[SD16] J. Seitz and M. Debes. *Kommunikationsnetze: Eine Umfassende Einfuehrung.* Ilmenauer Editionen, 2016.

[SDD13] D. Stanujkic, B. Dordevic, and M. Dordevic. Comparative Analysis of Some Prominent MCDM Methods: A Case of Ranking Serbian Banks. *Serbian Journal of Management*, 8:213 – 241, 2013.

[SdlOB16] M. I. Sanchez, A. de la Oliva, and C. J. Bernardos. Experimental Analysis of Connectivity Management in Mobile Operating Systems. *Computer Networks*, 94(C):41–61, January 2016.

[SE01] P. Srisuresh and K. Egevang. Traditional IP Network Address Translator (Traditional NAT). Request for Comments 3022, Internet Engineering Task Force, January 2001.

[SNW06] E. Stevens-Navarro and V.W.S. Wong. Comparison between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks. In *IEEE 63rd Vehicular Technology Conference*, volume 2, pages 947–951, 2006.

[SR11] N. Saxena and A. Roy. Novel Framework for Proactive Handover with Seamless Multimedia over Wlans. *IET Communications*, 5(9):1204–1212, June 2011.

[ssoI02] Telecommunication standardization sector of ITU. IP Packet Transfer and Availability Performance Parameters. ITU-T Y 1540 (Recommendation), November 2002.

[ssoI03] Telecommunication standardization sector of ITU. Network Performance Objectives for IP-Based Services. ITU-T Y 1541 (Recommendation), February 2003.

[Sta09] W. Stallings. *Data and Computer Communications.* Pearson Education; 8'th Edition, 2009.

[SV09] A. Sgora and D. D. Vergados. Handoff Prioritization and Decision Schemes in Wireless Cellular Networks: A Survey. *IEEE Communications Surveys Tutorials*, 11(4):57–77, Fourth 2009.

[Sys16] Cisco Systems. Cisco Visual Networking Index: Global Mobile Data Traffic Forcast Update 2015-2016 White Paper. Technical report, 2016.

[Tan02] A. S. Tanenbaum. *Computer Networks.* Pearson Education, 2002.

[TM89] E. Triantaphyllou and S. H. Mann. An Examination of the Effectiveness of Multi-Dimensional Decision-Making Methods: A Decision-making Paradox. *Decision Support Systems*, 5(3):303 – 312, 1989.

[TR05] M. Tuexen and M. Riegel. Mobile SCTP. Internet-Draft, January 2005.

[Var] A. Varga. *OMNeT++ User Manual version 4.3.*

[WB11] M. Wasserman and F. Baker. IPv6-to-IPv6 Network Prefix Translation. Request for Comments 6296, Internet Engineering Task Force, June 2011.

[WKG99] H. Wang, R. Katz, and J. Giese. Policy-eEabled Handoffs Across Heterogeneous Wireless Networks. In *Second IEEE Workshop on Mobile Computing Systems and Applications*, pages 51–60. Second IEEE Workshop on Mobile Computing Systems and Applications, 1999.

[Wol00] A. Wolisz. Wireless Internet Architectures: Selected Issues. In J. Wozniak and J. Konorski, editors, *Personal Wireless Communications*, pages 1–16. Kluwer, September 2000. Invited Paper.

[WS99] E. Wedlund and H. Schulzrinne. Mobility Support Using SIP. In *Proceedings of the 2Nd ACM International Workshop on Wireless Mobile Multimedia*, WOWMOM '99, pages 76–82, New York, NY, USA, 1999. ACM.

[YCK+10] H. Yokota, K. Chowdhury, R. Koodli, B. Patil, and F. Xia. Fast Handovers for Proxy Mobile IPv6. Request for Comments 5949, Internet Engineering Task Force, September 2010.

[YeN10] X. Yan, Y. Ahmet Şekercioğlu, and S. Narayanan. A Survey of Vertical Handover Decision Algorithms in Fourth Generation Heterogeneous Wireless Networks. *Computer Networks*, 54(11):1848 – 1863, 2010.

[YHI12] A. A. Yussuf, W. H. Hassan, and S. Issa. A Review of VHD Approaches in Next Generation Wireless Networks. In *Digital Information and Communication Technology and it's Applications (DICTAP), 2012 Second International Conference on*, pages 363–367, May 2012.

[Zha04] W. Zhang. Handover Decision Using Fuzzy Madm in Heterogeneous Networks. In *Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE*, volume 2, pages 653–658 Vol.2, March 2004.

[ZM06] F. Zhu and J. McNair. Multiservice Vertical Handoff Decision Algorithms. *EURASIP Journal on Wireless Communications and Networking*, 2006(2):1–13, April 2006.

List of Figures

List of Tables

Abbreviations

AHP	Analytical Hierarchical Process
AN	Anchor Node
AP	Access Point
AR	Access Router
BS	Base Station
BU	Binding Update
CIDR	Classless Inter-Domain Routing
CN	Correspondent Node
CoA	Care of Address
DiNAT	Dynamic index NAT
EID	End point Identifier
FA	Foreign Agent
FMIPv6	Fast MIPv6
GRA	Grey Relational Analysis
GTS	Global Translation Server
HA	Home Agent
HHO	Horizontal Handover
HIP	Host Identity Protocol
HMIPv6	Hierarchical MIPv6
HoA	Home Address

HTI	Handover Time Instants
InterNU	Inter NAT Update
IntraNU	Intra NAT Update
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
LCoA	Local CoA
LIN6	Location Independent Addressing for IPv6
LISP	Location Identifier Separation Protocol
LMA	Local Mobility Anchor
LTE	Long-Term Evolution
LTS	Local Translation Server
MADM	Multiple Attributes Decision Making
MAG	Mobile Access Gateway
MAP	Mobility Anchor Point
MCDM	Multiple Criteria Decision Making
MIFA	Mobile IP Fast Authentication
MIH	Media Independent Handover
MIPv4	Mobile IPv4
MIPv6	Mobile IPv6s
MMSP	Mobile Multimedia Streaming Protocol
MobileNAT	Mobile NAT
MOS	Mean Openion Score
MSCTP	Mobile Stream Control Transmission Protocol
mTCP	multi-path TCP
MTU	Message Transfter Unit
NAT	Network Address Translation

NAT64	IPv6-to-IPv4 Network Address Translation
NAT66	IPv6-to-IPv6 Network Address Translation
NID	Network Identifier
PAT	Port Address Translation
PMIPv6	Proxy MIPv6
QoMIFA	QoS-aware Mobile IP Fast Authentication Protocol
QoS	Quality of Service
RCoA	Regional CoA
REACH	Roaming-Enabled ArCHitecture
RLOC	Routing Locator
RSVP	Resource Reservation Protocol
SAW	Simple Additive Weighting
SEI	Session Entry Information
SIP	Session Initiation Protocol
SMT	Server Metrics
SNIR	Signal-to-Noise and Interference Ration
SR	Session Request
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
UMTS	Universal Mobile Telecommunications System
VHO	Vertical Handover
WiFi	Wireless Fidelity
WLAN	Wireless Local Area Network
WPM	Weighted Product Model
WSM	Weighted Sum Model