BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
FACULTY OF ELECTRICAL ENGINEERING AND INFORMATICS

Department of Networked Systems and Services

ADVANCED SCHEMES FOR EMERGING MOBILITY SCENARIOS IN THE ALL-IP WORLD

Ph.D. Dissertation
of

László Bokor

Supervisors:
Sándor Imre Sc.D.
Gábor Jeney Ph.D.

BUDAPEST, 2014

Declaration

I, undersigned László Bokor hereby certify that this dissertation, which I now submit for assessment on the programme of study leading to the award of Ph.D. is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Nyilatkozat

Alulírott Bokor László kijelentem, hogy ezt a doktori értekezést magam készítettem, és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Budapest, 2014. 06. 18.

…….
Bokor László

A dolgozat bírálata és a védésről készült jegyzőkönyv a későbbiekben a dékáni hivatalban elérhető.

Köszönetnyilvánítás

Mindenekelőtt szeretnék köszönetet mondani konzulenseimnek, Imre Sándornak és Jeney Gábornak. Útmutatásuk, hasznos tanácsaik és kritikáik nélkülözhetetlen segítséget nyújtottak kutatómunkámban és disszertációm elkészítése során egyaránt. Nekik tartozom azért is köszönettel, hogy színvonalas nemzetközi és hazai projektekben foglalkozhattam izgalmas kutatási-fejlesztési feladatokkal, és szakmai fejlődésemhez minden feltételt megteremtettek.

Különösen hálás vagyok hazai és külföldi szerzőtársaimnak és kollégáimnak az IP mobilitás területén együtt folytatott kutatásainkért, a közös munkáért és publikációkért, a rendkívül hasznos vitákért, és a konferenciák, projekt értekezletek során nem egyszer messzi országokban együtt szerzett élményekért.

Köszönet illeti a Mobil Kommunikáció és Kvantumtechnológiák Laboratórium, a Multimédia Hálózatok és Szolgáltatások Laboratórium, valamint a Mobil Innovációs Központ tagjait – közvetlen kollégáimat, akik magyarázataikkal, egy-egy hasznos mondattal, megjegyzéssel és tanáccsal nagyban könnyítették a munkámat.

Végül, ám korántsem utolsó sorban köszönettel tartozom egyetlen Pankámnak, szeretett szüleimnek és kedvenc húgocskámnak azért a biztos családi háttérért, melynél fontosabb feltétel nem létezett számomra doktori tanulmányaim során.

Kivonat

Napjainkban a telekommunikációs rendszerek különböző vezetékes és vezeték nélküli technológiák szinergikus egységévé formálódnak, melyekben Internet Protocol (IP) alapon futnak az integrált multimédia szolgáltatások. Az Internet egy teljesen átlátható és mindenütt jelenlévő multimédia kommunikációs rendszerré válik, melyben a felhasználók a távoli erőforrásokat bárhol és bármikor elérhetik. Az aktuális trendek és gyártói előrejelzések alapján kijelenthető, hogy a 2020-ig előttünk álló időszak a csomagkapcsolt mobil hálózatok forgalmának robbanásszerű növekedését fogja hozni. A várt forgalmi igények és felhasználói követelmények kielégítéséhez, a speciális használati esetek és komplex forgatókönyvek támogatásához a felhordó és maghálózati technológiáknak is fejlődniük kell. Ezen technológiákon belül kiemelkedő szereppel bírnak a mobilitás-kezelési protokollok és algoritmusok, melyek a jövő mobil Internének kulcsszereplői.

Disszertációmban új mobilitás-kezelési technikák kifejlesztésével, lokalizált mobilitás-menedzsment megoldások kidolgozásával, mikro-mobilitási tartományok tervezési kérdéseinek tárgyalásával és proaktív, rétegek közti (cross-layer) optimalizálásra támaszkodó hálózatváltási sémák bevezetésével céloztam meg a skálázhatóság növelését, az IP tartományok közötti észrevétlen mozgás támogatását, így végső soron az átviteli minőség és a felhasználói élmény javítását, valamint a felhasználók privát szférájának erősítését. Munkámat négy nagyobb témakörbe csoportosítottam.

A hagyományos makro-mobilitási protokollok skálázhatóságának és hálózatváltási teljesítményének növelését két megközelítést használva kívántam elérni. Egyrészről a Mobile IPv6 kiegészítése volt a célom, amit egy teljesen transzparens, decentralizált és a tartományon belül optimális utakat biztosító, IPv6 anycasting alapon működő mikro-mobilitási keretrendszer kidolgozásával értem el. Másrészről célul tűztem ki a mikro-mobilitás Host Identity Protocol (HIP) alapú jövő Internet rendszerekben történő támogatását is, ezért kifejlesztettem egy új, biztonságos jelzésdelegáción és címfordításon alapuló HIP mikro-mobilitási architektúrát és a hozzá tartozó, hálózatváltásokat kezelő protokollt.

Az IP világban történő mobilitás-kezelés során a felhasználó aktuális és mozgása során sokszor változó, követhető IP címe könnyedén átváltható precíz földrajzi pozícióadatokra. Második téziscsoportomban ezért a mikro-mobilitási megközelítések természetes lokátor-elrejtő képességét elemzem, és az általam kifejlesztett, felhasználók helyzetinformációnak védelmét támogató mikro-mobilitási tartománytervező algoritmusokat mutatom be, melyek segítségével a privát szféra védelmének egyre jelentősebb igényét már a mobil hálózatok tervezésekor figyelembe lehet venni.

A mozgó hálózatok (NEMO) mobilitás-kezelésének optimalizálását céloztam meg harmadik téziscsoportomban bemutatott két, eltérő megközelítést követő megoldásommal. Egyrészt a szabványos, IPv6 alapú hálózat-mobilitási protokollok hatékonyságának javításához hoztam létre egy egyedi, folyamatos hálózatmonitorozást és rétegek közti optimalizálást használó keretrendszert és speciális hálózatváltási sémát. Másrészt a mozgó hálózatok HIP rétegben való támogatásának hatékony biztosításához definiáltam egy új, HIP-alapú, jelzésdelegációra és mozgó randevúfunkcióra támaszkodó NEMO protokollt.

A jelenlegi erősen centralizált mobil Internet architektúrák nem skálázhatók az előre jelzett forgalmi növekménnyel, nem lesznek képesek kezelni a kihívásokat. Éppen ezért dolgoztam ki negyedik téziscsoportomban egy Host Identity Protocol alapú Ultra Flat Architecture keretrendszert, és végeztem el az architektúra szerves részét képző, proaktív, elosztott hálózatváltás-kezelő protokoll integrációját és teljesítmény-vizsgálatát. A javasolt megoldás előkészíti, és HIP-et használva végre is hajtja a hálózatváltásokat, eltűnteti az architektúrából a központosított IP horgonypontokat és a hálózati funkciókat a felhasználók közelébe helyezi, így segítve a skálázható mobil hálózati struktúrák kialakítását.

Abstract

Telecommunication industry predicts a huge mobile Internet traffic increase for the next decade with a series of emerging mobility scenarios and use-cases like network mobility for vehicles in Cooperative Intelligent Transportation Systems or scalable distributed mobility management for masses of mobile devices performing Machine to Machine communication. It seems to be technically challenging and prominently expensive to adapt current mobile network architectures and mobility management solutions to the increasing requirements. Core network technology must scale, novel protocols and design methodologies are needed to tackle the issues under limited revenue growth and increased user privacy. This work is to discuss advanced schemes and algorithms to support emerging mobility scenarios in future convergent distributed mobile Internet architectures.

In order to enhance legacy (macro)mobility management solutions by increasing their handover performance and scalability, I have followed two separate approaches. On the one hand I extended IPv6 with a novel, anycasting based micromobility extension for Mobile IPv6. Aiming at a transparent and distributed support of micromobility scenarios my goal was to propose a purely IPv6 based, and transparent micromobility framework. On the other hand I have exploited a candidate future Internet scheme built upon IP called the Host Identity Protocol, by designing and evaluating a novel HIP-based micromobility protocol naturally relying on the advanced, cryptographic ID/Loc separation scheme of HIP.

As mobility becomes one of the most unique characteristics of future's convergent architectures, more attention must be paid to the problems of location information leakage (i.e., location privacy issues of all-IP mobile communication caused by easy estimation possibilities from IP addresses to precise geographical positions of users), even at the earliest phases of design: at the network planning level. This motivated me to develop mobile network planning tools and algorithms that exploit inherent location privacy support of micromobility protocols.

For network mobility (NEMO) scenarios several improvements exist to overcome the limitations of the already standardized NEMO Basic Support protocol. However there are several extensions of NEMO BS, the searching for further optimization possibilities and novel solutions has not stopped. In order to enhance current NEMO schemes, I have followed two approaches. On the one hand I improved standard IPv6-based network mobility by forming a framework based on a special handover solution using location information support, cross-layer optimization and continuous network discovery. On the other hand I have further extended the Host Identity layer by developing and evaluating a novel, HIP-based NEMO protocol.

It is highly expected that due to their centralized (anchor-based) design, mobile Internet architectures currently being under deployment or standardization will not scale particularly well to efficiently handle the challenges. In order to overcome these issues, I have developed a Host Identity Protocol based system framework for the Ultra Flat Architecture, and also designed and evaluated a proactive, distributed handover preparation and execution protocol for this framework.

By covering the above emerging scenarios with optimized schemes and advanced algorithms for the all-IP world in my dissertation I was able to improve the performance of current solutions and thus increase the quality of mobile applications and the level of mobile user experience in general.

List of Abbreviations

3GPP	3rd Generation Partnership Project
802.21 MIH	802.21 Media Independent Handovers
AA	Anycast Address
AAA	Authentication Authorization Accounting
ABMF	Anycast Based Micromobility Framework
aCoA	Anycast Care-of-Address
AGM	Anycast Group Membership
AI	Anycast Initiator
AM-LSA	Anycast Membership LSA
ANDSF	Access Network Discovery and Selection Function
ANP	Access Network Predictor
AOSPF	Anycast Open Shortest Path First
AP	Anycast Prefix
AP	Access Point
AR	Anycast Router
AR	Access Router
ARD	Anycast Receiver Discovery
ARI	Anycast Route Information
ARIP	Anycast Routing Information Protocol
AS	Anycast Subnet
B2BUA	Back-to-Back User Agent
BEET	Bound-End-to-End-Tunnel
BEX	Base Exchange
BID	Binding Update
BOSS	On Board Wireless Secured Video Surveillance
BSSID	Basic Service Set Identifier
BU	Binding Update
BW	BandWidth
C UFA GW	Candidate UFA GW
CAPEX	Capital Expenditures
CAR	Correspondent Anycast Responder
CDN	Content Delivery Networking
CDTR	Context Transfer Data Reply
CIP	Cellular IP
C-ITS	Cooperative Intelligent Transportation System
CN	Correspondent Node
CoA	Care-of-Address
CPH	Control Plane Header
CR	Cell Border Crossing Rate
CSN	Connectivity service network
CTD	Context Transfer Data
CXTP	Context Transfer Protocol
D-H	Diffie–Hellman
DHCP	Dynamic Host Configuration Protocol
DIMA	Distributed IP Mobility Approach
DMM	Distributed Mobility Management
DNS	Domain Name System
DoS	Denial-of-Service
DSRC	Dedicated short-range communications
EPC	Evolved Packet Core
ESP	Encapsulating Security Payload
GGSN	Gateway GPRS Support Node
GMA	Gateway Mobility Agent
GNSS	Global Navigation Satellite System
GPRS	General Packet Radio Service
GPS	Geographical Positioning System
GREAL	Greedy LA Forming Algorithm
GSM	Global System for Mobile Communications

GTP	GPRS Tunneling Protocol
GUI	Graphical User Interface
HA	Home Agent
HAWAII	Handoff-aware Wireless Access Internet Infrastructure
HDTV	High Definition Television
HI	Host Identifier
Hi^3	Host Identity Indirection Infrastructure
HIP	Host Identity Protocol
HIP RS	HIP Rendezvous Service
HIP SM	HIP State Machine
HISM	Host Identity Specific Multicast
HIT	Host Identity Tag
HM	Handover Manager
HMIPv6	Hierarchical MIPv6
HSPA	Hight Speed Packet Access
HSS	Home Subscriber Server
HTTP	HyperText Transfer Protocol
ID/Loc	Identifier/Locator
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IKEv2	Internet Key Exchange version 2
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ITU-T	International Telecommunication - Union Telecommunication Sector
L1/L2/L3	Layer1/Layer2/Layer3 of TCP/IP
LA	Location Area
LAFA	Location Area Forming Algorithm
LFN	Local Fixed Node
LIN6	Location Independent Networking for IPv6
LIPA	Local IP Access
LMN	Local Mobile Node
LP	Location Privacy
LRVS	Local Rendezvous Server
LSA	Link State Advertisements
LSI	Local Scope Identifier
LTE/LTE-A	Long Term Evolution/Long Term Evolution-Advanced
M2M	Machine-to-Machine
MA	Mapping Agent
MAC	Medium Access Control
MAP	Mobile Anchor Point
MCoA	Multiple Care-of Addresses Registration
MEVICO	Mobile Networks Evolution for Individual Communications Experience
MIHF	Media-Independent Handover Function
MIHU	MIH User
MIIS	Media Independent Information Service
MIPv4	Mobile IPv4
MIPv6	Mobile IPv6
MitM	Man-in-the-Middle
MN	Mobile Node
MNN	Mobile Network Nodes
MOSPF	Multicast Anycast Open Shortest Path First
MPLS	Multiprotocol Label Switching
MR	Mobile Router
mRVS	Mobile Rendezvous Point
ND	Neighbor Discovery
NEMO	Network Mobility
NEMO BS	NEMO Basic Support
OPEX	Operation Expenditure

P2P	Peer-to-Peer
PAI	Paging Area Identifier
PAN	Personal Area Network
PA-SABLAF	Privacy Aware SABLAF
PCC	Policy and Charging Control
PDA	Personal Digital Assistant
PDN GW or PGW	Packet Data Network Gateway
PHY	Physical
PIA-SM	Anycast Protocol Independent Anycast - Sparse Mode
PIM-SM	Protocol Independent Multicast - Sparse Mode
PoA	Point of Access
PoS	Point of Service
PRD	Paging Registration Database
PUA	Peer Unicast Address
QoE	Quality of Experience
QoS	Quality of Service
RA	Router Advertisement
RegReg6	Regional Registrations
RFC	Requests for Comments
RFID	Radio Frequency Identification
RIPng	Routing Information Protocol Next Generation
RP	Rendezvous Point
RP 3/5	Reference Point 3/5
RSSI	Receive Signal Strength Indicator
RTT	Round Trip Time
RVS	Rendezvous Server
S UFA GW	Source UFA GW
SABAS	Simulated Annealing Based Anycast Subnet forming
SABLAF	Simulated Annealing Based Location Area Forming Algorithm
SAE	System Architecture Evolution
SAP	Service Announcement Packet
SHV	Super Hi-Vision
SIP	Session Initiation Protocol
SIPTO	Selected IP Traffic Offload
SMS	Short Messaging Service
SNR	Signal-to-Noise Ratio
SPI	Security Parameter Index
SPINAT	Security Parameter Index Network Address Translation
SSID	Service Set Identifider
T UFA GW	Target UFA GW
TB-LAD	Traffic-Based Static Location Area Design
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
UDP	User Datagram Protocol
UFA	Ultra Flat Architecture
UFA GW	UFA Gateway
UFA-HIP	HIP-based Ultra Flat Architecture
ULP	User Location Privacy
UMTS	Universal Mobile Telecommunications System
USB	Universal Serial Bus
VAG	Virtual Anycast Group
VAN	Vehicle Area Network
VMN	Visiting Mobile Node
VoD	Video on Demand
VoIP	Voice over IP
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WR	Weighted Rate
XML	Extensible Markup Language

Contents

List of Figures and Tables

Chapter 1
Introduction

Telecommunication systems are converging into a synergistic union of different wired and wireless technologies, where integrated, multimedia services are provided on a universal IP-based infrastructure [J1], [C7]. Besides the evolution of wireless networks toward heterogeneous all-IP mobile communication architectures, end-user terminals are also becoming more and more powerful implementing extremely large variety of functions from making voice and video calls through social networking and sharing multimedia till exploiting the advantages of geographic positioning solutions [C8]. The Internet itself is turning into a fully pervasive and ubiquitous communication system in which users are expected to be able to use remote resources anytime and anywhere. This evolution recently made mobile Internet a reality for both users and operators thanks to the success of novel, extremely practical smartphones, portable computers with easy-to-use 3G USB modems and attractive business models. Based on actual trends in telecommunications, vendors prognosticate that mobile networks will suffer an immense traffic explosion in the packet switched domain up to year 2020 [1]–[4]. In order to accommodate current systems to the anticipated traffic demands and user requirements, technologies applied in the access, backhaul and core networks must become appropriate to advanced use cases and scenarios. Within these technologies, mobility management protocols and schemes play an essential role when it comes to future mobile Internet architectures [J9].

Legacy IP mobility management solutions like Mobile IPv4/IPv6 [5], [6] provide transparent session continuity and global handover management for heterogeneous all-IP mobile communication architectures but could suffer from several well known problems (increased delay, packet loss, and signaling) that have led to the distinction of macro- and micromobility scenarios. Macromobility focuses on mobility management between distant wireless domains and across the Internet [5]–[8], [C16], [J4], while protocols designed for micromobility scenarios [9]–[11] reduce the number of network elements that process the signaling information by restricting the propagation of such datagrams to a smaller set of nodes and manage movement inside a specific wireless domain locally. Due to their performance and scalability during handovers within localized areas, optimization, development and integration of micromobility schemes are research topics that live their renaissance nowdays. The optimal design of micromobility domains is also an open issue when deploying these protocols in next generation mobile environments.

Trends clearly show that IP-based mobile and wireless networks will not only support mobility for the widest range of single end terminals, but even for Personal Area Networks (PANs), Vehicle Area Networks (VANs) [12], complex groups of nodes in Intelligent Transportation Systems (ITSs) and Cooperative ITS (C-ITS) architectures [13], [C5], [C10], [C15] complete networks of RFID (Radio Frequency Identification) devices and sensors, and various mobile ad hoc networks [14]. It means that not only single mobile entities with permanent Internet connectivity have to be managed, but also entire mobile networks (i.e., NEMOs) need to be maintained as a whole. The currently standardized NEMO protocol [15] only offers basic solution for this complex problem, thus leaving space for researches on further enhancement and optimization.

The growing number of mobile users, the increasing traffic volume, the complexity of mobility scenarios, and the development of new and innovative IP-based applications require network architectures and protocols able to deliver all kind of traffic demands seamlessly assuring high end-to-end quality of service. However, the strongly centralized nature of current and planned mobile Internet standards (e.g., the ones maintained by the IETF or by the

collaboration of 3GPP) prevents cost effective system scaling for the novel traffic demands. Micromobility protocols try to ease the above issues, but doesn't find the root of the problem. Aiming to solve the burning questions of scalability from an architectural point of view, distributed [16],[J11] and flat [17] mobile architectures with enhanced, proactive and cross-layer optimized techniques (e.g., [C23], [C30]) are gaining more and more attention today.

However IPv6 shows word-wide proliferation and will play an essential role in future communications, it is also anticipated in next generation mobile architectures that IP addresses will not continue to remain both locators (for packet routing) and identifiers (for referring to a host or session): the semantically overloaded nature of the Internet Protocol will be obviated by identifier/locator (ID/Loc) separation schemes [18], [19]. The Host Identity Protocol (HIP) family [20]–[23] is one of the most promising, extendable and flexible ID/Loc separation techniques, which guided me to develop both HIP and pure IPv6 based solutions for the identified problems.

1.1. Research Objectives and Thesis Structure

The above introduced trends and use-cases pose serious challenges to existing mobile Internet architectures and require special support to efficiently cope with the raised problems and questions. My essential aim was to develop advanced protocols and schemes supporting these emerging mobility scenarios of the all-IP world. By investigating new mobility management techniques, localized mobility solutions, micromobility domain planning algorithms and proactive, cross-layer optimized handover mechanisms, I could also ensure scalability, seamless handover, enhanced network design, and eventually better Quality of Service (QoS), Quality of Experience (QoE) and increased user privacy. Regarding to the previously summarized broad research areas I have grouped my researches into four main topics:

1. In order to enhance macromobility management solutions by increasing their handover performance and scalability, I have followed two separate approaches. On the one hand I was induced to investigate possibilities to enhance the Internet Protocol and design a novel micromobility extension for Mobile IPv6 (Thesis I.1 and I.2). Aiming at a transparent and distributed support of micromobility scenarios my goal was to propose a purely IPv6-based, and transparent micromobility framework, which doesn't require additional network entities, provides highly decentralized operation, and ensures optimal routes inside the domains without introducing extra signaling load on the wireless interface. In order to support deployment by keeping the scalability and efficiently controlling the size of the micromobility routing domain in the network design phase, the development of a special subnet optimization algorithm for my framework was also an objective within this approach. On the other hand I have decided to exploit a candidate future Internet scheme built upon IP called the Host Identity Procotol, by designing and evaluating a novel HIP-based micromobility protocol (Thesis I.3) naturally relying on the advanced, cryptographic ID/Loc separation scheme of HIP.

   Thesis I.1: A built-in IPv6 micromobility management scheme based on anycasting (ABMF) is introduced in Section 2.1.2.

   Thesis I.2: A special anycast subnet forming algorithm is developed and evaluated in an improved mobility simulator in Section 2.1.3.

   Thesis I.3: A localized mobility management extension of Host Identity Protocol (μHIP) is presented in Section 2.2.2. An accurate HIP simulation environment is developed and used for accurate modeling and evaluation of μHIP in Section 2.2.3.

2. As mobility becomes one of the most unique characteristics of future's convergent architectures, more attention must be paid to the problems of location information leakage (i.e., location privacy issues of all-IP mobile communication caused by easy estimation possibilities from IP addresses to precise geographical positions of users), even at the earliest phases of design: at the network planning level. This motivated me to develop mobile network planning tools and algorithms that exploit inherent location privacy support of micromobility protocols (Thesis II.1, II.2, II.3, and II.4). Existing network planning algorithms (e.g., [24]–[28]) are mainly focusing on the trade-off between the paging cost and the registration cost and – to the best of my knowledge – none have introduced privacy awareness in network planning methodologies before my work.

   Thesis II.1: A location privacy policy model for micromobility domain planning with an appropriate algorithm (PA-SABLAF) is discussed in Section 3.1.1.
   Thesis II.2: Performance of PA-SABLAF is evaluated with the help of a proprietary location privacy metric in Section 3.1.2.
   Thesis II.3: A PA-SABLAF variant using uncertainty-based location privacy metric is presented end evaluated in Section 3.2.2.2.
   Thesis II.4: A PA-SABLAF variant using traceability-based location privacy metric is presented end evaluated in Section 3.2.2.2.

3. For network mobility scenarios several improvements exist to overcome the limitations of the already standardized NEMO Basic Support protocol [15]. NEMO BS operates in the IP layer and inherits the benefits of Mobile IPv6 [6] by extending the binding mechanism of the ancestor, but keeps all the problems of the main approach such as protocol overhead, inefficient routing, security and lack of multihoming support. All of these issues are under examination at the IETF, but this work has not been completed yet. However, there are several extensions of NEMO BS in order to allow multihoming and nested mobile networking [29], [30], and ongoing researches are trying to deal with the route optimization [31]–[33], security problems [34], [35], and handover optimization [36]–[38]. Despite the fact that several novel real-life demonstrations [39] and testbeds [40] started to prove the feasibility and usability of NEMO BS and its extensions, the searching for further optimization possibilities and novel solutions like [41] has not stopped. In order to enhance current NEMO schemes, I have followed two approaches. On the one hand I was aiming at improving standard IPv6-based network mobility by forming a framework based on a special handover solution (Thesis III.1 and III.2) using cross-layer optimization and continuous network discovery. On the other hand my goal was to extend the Host Identity layer by developing and evaluating a novel, HIP-based NEMO protocol (Thesis III.3).

   Thesis III.1: A location information aided predictive mobility management framework for multihomed NEMO BS configurations is introduced in Section 4.1.2.
   Thesis III.2: The prediction accuracy of the proposed solution is analyzed using a probabilistic model in Section 4.1.3.
   Thesis III.3: A Host Identity Protocol based network mobility solution (HIP-NEMO) is presented in Sections 4.2.2.1, 4.2.2.2, and 4.2.2.3. The performance evaluation of HIP-NEMO is provided based on extensive simulations built on complex protocol models in Section 4.2.2.4.

4. It is highly expected that due to their centralized (anchor-based) design, mobile Internet architectures currently being under deployment or standardization will not scale particularly well to efficiently handle the challenges [42], [J9]. To enhance scalability of

mobile Internet architectures and support distributed mobility management scenarios with decentralized, proactive, self-configuring and self-optimizing network structures, the Ultra Flat Architecture (UFA) was proposed as one of the first solutions [17], [43]. The main characteristic of this proposal is that the execution of handovers is managed by the network via the Session Initiation Protocol (SIP) [44]. Even though SIP is a very powerful signaling solution for UFA, it is not applicable for non-SIP (i.e., legacy Internet) applications and the published SIP-based UFA scheme also does not comply with ITU-T's recommendation of requirements for ID/Loc separation in future networks [18]. In order to overcome these issues, my research objective was to develop a Host Identity Protocol based system framework for the Ultra Flat Architecture (Thesis IV.1), and also to design and evaluate a proactive, distributed handover preparation and execution protocol for this framework, supporting complete elimination of centralized IP anchors between Point of Access (PoA) nodes and correspondent nodes, and placing network functions at the edge of the transit and access networks (Thesis IV.2 and IV.3).

Thesis IV.1: A Host Identity Protocol based system framework for the Ultra Flat Architecture (UFA-HIP) is proposed in Section 5.1.2.

Thesis IV.2: A proactive, 802.21 MIH and HIP-based handover initiation, preparation, execution and completion protocol for UFA-HIP is presented in Section 5.2.2.

Thesis IV.3: The performance of the proposed UFA-HIP handover protocol is evaluated in Section 5.2.3.

1.2. Research Methodology

In my Thesis I have relied on two classical research approaches: analytical considerations and simulation studies. During the development phase of novel protocols, schemes or algorithms for the identified problems of emerging mobility scenarios, analytical considerations could not be ignored. My work on special network planning solutions in Thesis groups I and II is based on graph models, cost structures, and theory of algorithms (i.e., simulated annealing), while the analysis of my special NEMO optimization framework in Thesis group III relied on probability theory.

My proposed schemes were implemented in two different simulators. On the one hand I modified and extended an existing, proprietary Java-based mobility simulator [45], [J3], producing realistic cell boundary crossing (i.e., inter-cell movement rate) values and incoming call database in the particular (micro)mobility system under evaluation in Thesis group I and II. This simulator provided a realistic representation of the mobility patterns and was prepared to execute the different algorithm variants over an initial domain structure. On the other hand I have modified and extended an existing C++ model package for a general purpose open-source, component-based, discreet event simulation environment called OMNeT++ [46]. Thesis groups I, III and IV rely on the extensive evaluations performed with the help of my contributions to this powerful environment [46], [C17].

I have strongly relied on statistics and probability theory also within my simulation analysis when handling large amount of measurement data came into picture.

Chapter 2
Micromobility Management Protocols

Rapid evolution of wireless networking has provided wide-scale of different wireless access technologies (e.g., 802.11a/b/g, DSRC, 3G UMTS, LTE, LTE-A, WiMAX, etc.) with complementary characteristics and motivation of operators to integrate them in a supplementary and overlapping manner. To provide ubiquitous mobility between these technologies, Internet Protocol v4 and v6 emerged as the common technology platform [J5], [B6] which is capable of connecting the various wired and wireless networks. Although macromobility management protocols (e.g., Mobile IPv4 [5] and Mobile IPv6 [6]) are capable of handling global mobility of users, they introduce low scalability, significant signaling overhead, and increased delay and packet loss when mobile terminals change their Internet point of attachment (PoA) frequently within geographically small areas (i.e., micromobility domains) [47]. In order to overcome these performance deficiencies, several approaches attempt to extend IP level global macromobility mechanisms: micromobility methods (e.g., [9]–[11], [48]) offer faster and more seamless handover management while also reduce load on central mobility anchor points and (e.g., [49]) enable more scalable operation and resource utilization. However these approaches usually suffer from lack of robustness, inefficient handling of intra-domain traffic and added complexity, furthermore they often require employing of new protocol stacks, and in general do not offer optimal performance in several scenarios.

In order to enhance macromobility solutions by increasing their transparency, handover performance and scalability, I have followed two separate approaches. On the one hand I have investigated possibilities to enhance the Internet Protocol and designed a purely IPv6-based micromobility extension for Mobile IPv6 (Thesis I.1 and I.2 in Section 2). On the other hand I have exploited a candidate future Internet scheme built upon IP called the Host Identity Procotol (HIP) [20]–[23] and designed a HIP-based micromobility protocol (Thesis I.3 Section 2.2).

2.1. Built-in IPv6 Micromobility Management based on Anycasting

In my IPv6-based scheme the main goal was to rely on the characteristics and latest results of the IPv6 anycasting, and such providing a built-in and transparent solution for micromobility management.

Thesis I.1. [C1],[C2],[C3],[B1] *I have proposed an anycast based micromobility framework (ABMF), which provides completely distributed, highly decentralized operation and optimal routes inside the micromobility domains without introducing extra signaling load on the wireless interface.*

2.1.1 Overview of IPv6 Anycasting

Anycasting is a group communication scheme which was introduced originally in RFC 1546 [50]. Anycasting separates service identifiers from physical interfaces, enabling a service to act as a logical entity of the network. Several promising practical application can be imagined based on this characteristics. The most popularly known application of anycast technology is helping the communicating nodes in selection of service providing servers. In

this approach the client host can choose one of many functionally identical servers. As a result, load distribution and balancing can be achieved between the multiple servers when we use a feasible anycast routing protocol, where anycast requests are fairly forwarded. An excellent survey of the IPv6 anycast characteristics and applications can be found in [51], [52], where the authors describe many advantages and possible applications of anycasting.

The anycasting paradigm was adopted in IPv6 as one of its basic and explicitly included services [53]. When an IPv6 node sends a packet to an anycast address, the network (based on underlying routing algorithms) will deliver the packet to at least one and preferably only one of the competent hosts thus establishing one-to-one-of-many communication. In this matter IPv6 anycasting is considered as a group communication scheme, where the group of nodes is represented by an anycast address and anycast routing algorithms are dedicated always to find the most appropriate destination for an anycast packet. The "appropriateness" is measured by the metric of the routing protocol. In IPv6 the anycast addresses cannot be distinguished from the unicast addresses, they share the same address space. Therefore the beginning part of IPv6 anycast addresses is the network prefix: the longest P prefix identifies the topological region in which the anycast group membership must be handled as a separate host entry of the routing system. Outside this region anycast addresses of that membership can be aggregated. Existing drafts categorize IPv6 anycast based on the length of P [54]. On the one hand Global Anycasting should be taken into consideration, where the value of the P prefix is zero, making aggregation impossible and leading to serious scalability problems: individually stored anycast entries easily could cause explosion of routing tables if anycasting gets widely used. On the other hand Subnet Anycasting should be considered when anycast packets can reach the last hop router by normal unicast routing, and the current Anycast Responder is determined by the last hop router (e.g. based on Neighbor Discovery). Regional Scoped Anycasting [55] is a natural outgrowth of Subnet Anycasting: the anycast subnet may contain not only one router (i.e. the last hop router) but more, creating a controlled anycast subnet (or region) by restricting the advertisement of anycast routing information (Fig. 1).

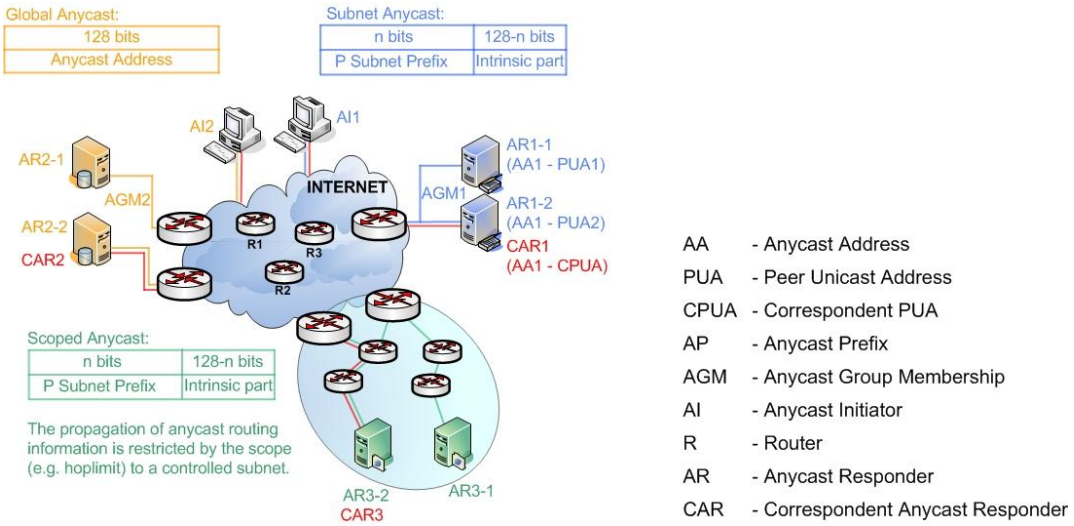


Figure 1: Terminology of IPv6 anycasting

Anycast routing protocols working in the subnet (i.e. scope-controlled region) should take care of managing the anycast membership and exchanging the anycast routing information. The current IPv6 standards do not define the anycast routing protocol, although the routing is one of the most important elements of network-layer anycasting. Beyond the lack of standards, there is quite small amount of literature about practical IPv6 anycasting.

However the existing drafts are quite prosperous [56], [57], there are still challenges to be solved.

V. Park and J. Macker proposed anycast extensions of link-state routing algorithm and distance-vector routing algorithm in [58] and evaluated in [59]. D. Xuan and others proposed and compared several routing algorithms for anycast [60]. Eunsoo Shim proposed an application load sensitive anycast routing method (ALSAR) and analyzed the existing routing algorithms in his PhD thesis [61]. S. Doi and others summarized the problems and possible solutions regarded the current specifications for IPv6 anycasting and proposed an anycast routing architecture based on seed nodes, gradual deployment and the similarities to multicasting [51]. Based on their work S. Doi and others together with S. Matsunaga and others designed and implemented three IPv6 anycast routing protocols (AOSPF, ARIP, PIA-SM) based on existing multicast protocols (MOSPF, RIPng, PIM-SM) [56], [57]. The area of secure and reliable anycast group membership management protocol is also being investigated (e.g., [62]), as well as the problems coming from the stateless nature of anycasting [51]. Due to promising achievements in the area of IPv6 anycasting, the restrictions introduced in the early IPv6 standards (RFC 3513 [63]) are now removed (RFC 4291 [53]), proving that the IPv6 community will sooner or later come up with a standardized solution.

2.1.2 IPv6 Anycast based Micromobility Framework (ABMF)

In the proposed IPv6 mobility management framework the anycast addresses are identifying the mobile nodes (MNs) entering a micromobility domain. In the micromobility domains the registering and the membership management of the mobile anycast nodes is done by anycast group membership management protocols like [62] or [64]. The location- and handover management of mobile nodes within a given micromobility domain (i.e., intra-domain communication of a given anycast subnet) is based on the underlying anycast routing protocol (e.g., [56], [57], [65]). Inter-domain handovers are managed with the well-known Mobile IPv6 macromobility protocol.

In ABMF, when a mobile node enters a micromobility domain, the Care-of-Address (CoA) obtained is a unique anycast address (aCoA), thus an anycast address identifies a single mobile node. Therefore the packets sent to the aCoA of the mobile terminal have no chance of reaching another mobile node, since in this sense the anycast addresses assigned to the mobile nodes are unique. The assigned anycast address has a validity area or region – an Anycast Subnet (AS) defined by the P prefix and the scope – where the anycast address might be located. As a result the mobile node in the validity area of the anycast address can move without being forced to change its anycast Care-of-Address. The mobile node with a unique anycast Care-of-Address matches the Correspondent Anycast Responder (CAR) in anycasting terminology [54]. In my scheme the validity area determined by the length of the P prefix of the anycast address equals a micromobility domain. As a result the movements within the micromobility domain (i.e., anycast subnet) are handled locally decreasing the signalling overhead of MIPv6 as the corresponding macromobility protocol.

Within the micromobility domain the use of anycast address as an identifier for the mobile terminals helps to simplify the routing and handover management by applying routing metrics. As a result the movements of the mobile nodes can be characterized by the change of the routing metrics in the anycast routing tables; no new routing entries are needed when moving.

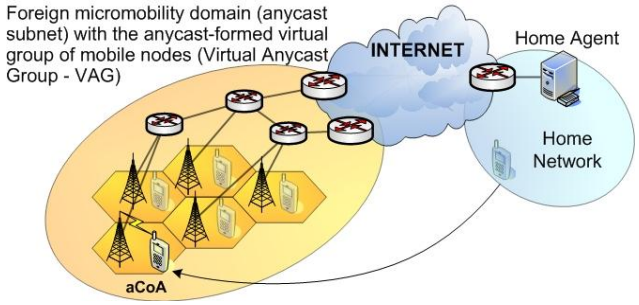


Figure 2: IPv6 Anycast-based Mobility Framework (ABMF)

The mobile node after entering a micromobility domain and getting an anycast CoA becomes a member of a Virtual Anycast Group (VAG). The VAG size depends on the size of the micromobility area (or anycast subnet) since the anycast address is valid in the whole micromobility domain. The members of the VAG are the virtual (possible) locations of the mobile node (Fig. 2). However the mobile node's actual position is the only one that has a valid routing entry. The Virtual Anycast Group equals the Anycast Group Membership (AGM) while the virtual copies of the mobile node match the Anycast Responders according to [52]. The movement of the mobile node equals the joining of a new Anycast Responder (at the new location of the mobile node) and the quitting of an old Anycast Responder (from the previous site). The underlying anycast routing algorithms are supposed to find out the appropriate destination for a packet destined to a VAG member.

2.1.2.1 ARIP operation in ABMF

One of the most important infrastructural basics regarding any anycast based application is the underlying routing protocol. In order to show how my ABMF proposal would work in practice I introduce the solution's four main scenarios using Anycast Routing Information Protocol (ARIP) [57].

In the first scenario the mobile terminal leaves its current domain (e.g., its Home Network) and enters (1) an other local administrative mobility domain (a new micromobility domain) as seen in Fig. 3 case the mobile node first of all obtains (2) – with the help of address autoconfiguration method on receiving a Router Advertisement – a unique anycast address that is valid in the whole area due to the properly set *P* prefix of the anycast address. As a result the source address can be a unique anycast address since the source of a packet can be identified unequivocally. After getting the unique anycast Care-of-Address, the mobile node has to build the binding towards its Home Agent therefore a Binding procedure (3) is started by sending a Binding Update message. Next the mobile terminal has to initiate its Anycast Membership in the Virtual Anycast Group (VAG) of the new micromobility domain by having its anycast CoA. This can be done with the help of an ARD (Anycast Receiver Discovery) Report message (4). On receiving an ARD report message the access router creates an ARI (Anycast Route Information) message (5) and sends it towards it adjacent routers that insert the received information into their routing table associated with the output interface information. As a result each router in the new micromobility domain has an entry in their routing table how to reach the mobile terminal. Since each outing entry has a timeout period thus the mobile node should send ARD Report message periodically to maintain its routing entry. The updating time of the routing entry should be defined according to the refresh interval of the routing entries.

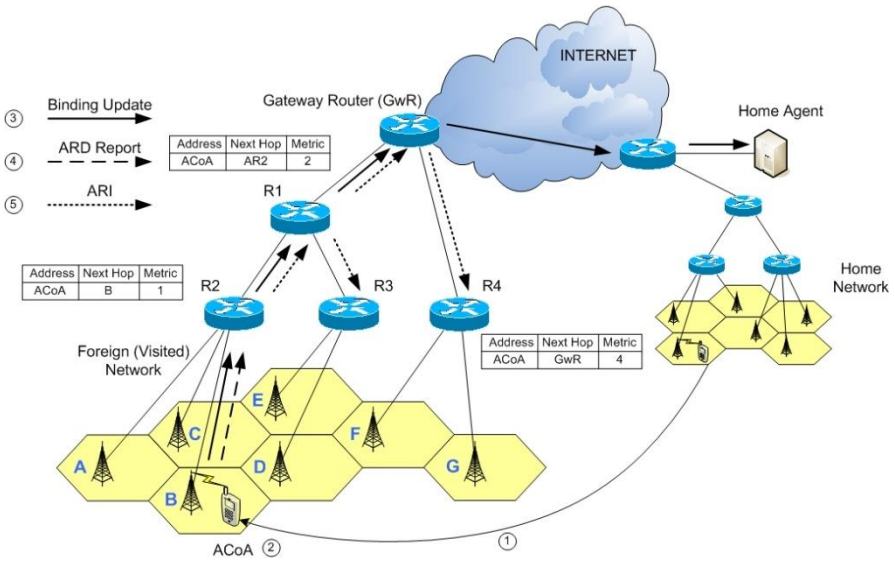


Figure 3: Entering a foreign micromobility domain

In the second scenario (Fig. 4) the mobile node moves (1) while sending data packets ("ready state") toward the Correspondent Node (i.e., the Anycast Initiator). As the mobile terminal is attached to the new access router, the new router notices that packets with the anycast address in the source address field are being sent over one of its interfaces (2) (the access router checks the direction where it receives the anycast-sourced packets). According to the anycast routing protocol the access router has an entry in its routing table regarded this source anycast address. Therefore the router modifies the entry regarded the anycast address of the mobile node so that the new entry forwards the packets towards their new destination (the interface from which it has received the packet with the anycast address in the source address field), the actual location of the mobile terminal. The access router also has to initiate an anycast routing information exchange by sending an ARI message (3). In this case the ARI message should only propagate to the previous router since the rest of the path remains unchanged. The previous access router can be reached easily since the router entry before the update points towards the previous router.

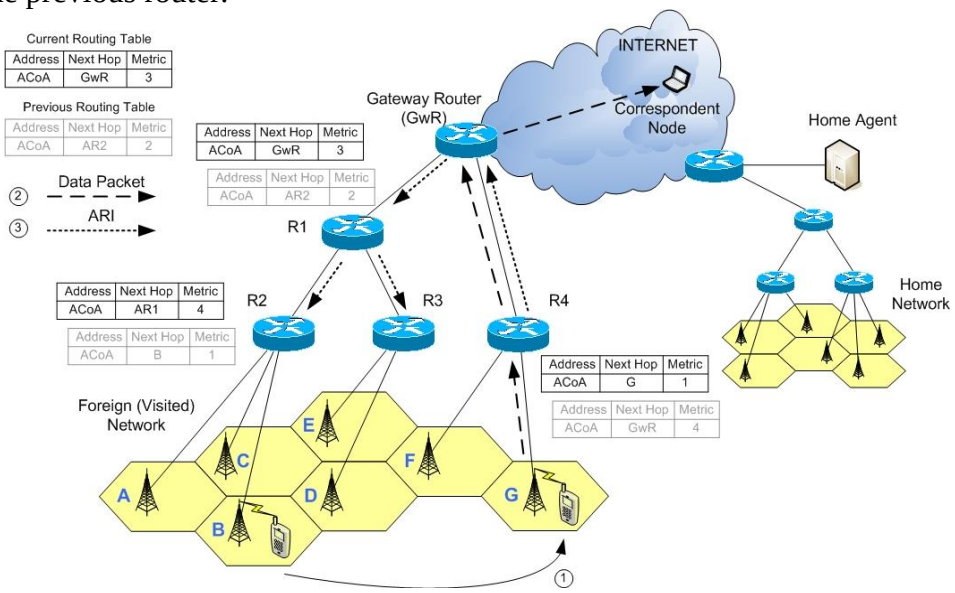


Figure 4: Moving in a given micromobility domain

In the third scenario the mobile node changes its point of attachment in a stand-by state (the mobile is attached to the network and involved in mobility management, but there is no data transmission). As the mobile node notices the change of the access router (based on timers, router advertisement or some kind of lower layer trigger) the mobile terminal informs the network of its current location by sending an ARD Report message to the new router. As it has been shown in the previous scenario the same applies here since the new access router is responsible to spread the routing information throughout the micromobility domain with the help of ARI message exchange.

In the fourth scenario the mobile node is in idle state (the MN is not reachable, it is not attached to the network) while moving around the micromobility domain. As a result if a packet arrives to the mobile terminal's anycast address the mobile node has to be found therefore a special ARD Query is sent throughout the network to find the actual location of the mobile terminal. This implements a simple paging mechanism aiming to give the protocol a reactive attitude and decrease the signaling overhead over the radio link. In the special ARD Query the anycast Care-of-Address of the mobile node should be inserted as well, since only the mobile node with the unique anycast address should reply to the ARD Query message. As the mobile node answers to the ARD Query with an ARD Report message thus the routing information can be distributed in the micromobility domain.

Of course a more complex paging scheme could also be implemented, but it will decrease the transparency of the proposed scheme. On the one hand when a mobile node is active or participating in the routing information exchange, the maintaining of paging cache can be done by the active signalling or by the ongoing communication's packets. On the other hand when the mobile node is in idle mode the balance should be found to keep the paging cache up-to-date and to keep the signalling overhead low. In my proposed framework the Mobile IPv6 Binding Update message can be used to keep the paging cache up-to-date. The mobile node should send periodic BU messages towards its Home Agent to refresh the binding. The routers – on the path of the BU towards the Home agent – in the paging area refresh the paging cache on the arrival of a BU message. In ABMF the paging would follow a distributed approach since the paging cache is distributed in the anycast routers of the micromobility domain. Therefore the risk of a single point of failure in the paging cache can be reduced [49].

2.1.2.2 AOSPF operation in ABMF

Not only ARIP can be used in my ABMF proposal, but any other IPv6 anycast routing scheme can be applied. To highlight this transparency I have selected the Anycast Extension to OSPFv3 [57], [65] as an alternative underlying routing infrastructure. The AOSPF routing protocol and the AGM management works closely together in order to maintain the routing information flow similarly to the ARIP case. The main differences are depicted below (Fig. 5).

1) The topology of the anycast routers is created with the help of the Link State Advertisements (LSA). This step is same as in case of the standard OSPFv3 [66].
2) As indicated previously, the movements of the mobile node can also be represented by the VAG membership information exchanging. This is done with the help of the Anycast Receiver Discovery (ARD) queries and reports (or can be done by other, more secure anycast group membership management protocols).
3) The anycast router upon receiving an ARD report creates an Anycast Membership LSA (AM-LSA) packet. The AM-LSA then is sent to the adjacent anycast routers.
4) The anycast router when receiving an AM-LSA message checks whether the received anycast address has already been stored in the routing table. In case it is a new entry (e.g., a new mobile node arrives to the micro mobility domain), the

anycast router simply registers it, and then forwards the entry to the adjacent routers. Otherwise (when there was a previous routing entry) the appropriateness of the newly arrived AM-LSA is evaluated. Note that the ABMF framework simplifies the evaluation phase, since the arrival of a new AM-LSA message would mean that the mobile node has moved. Therefore the latest AM-LSA message contains the latest – and most up-to-date – information about a single mobile node. The propagation of the AM-LSA messages can also be limited since when the AM-LSA message does not generate any change in the routing table (the AM-LSA reached a crossover node) then the router should not forward the AM-LSA message towards its adjacent routers.

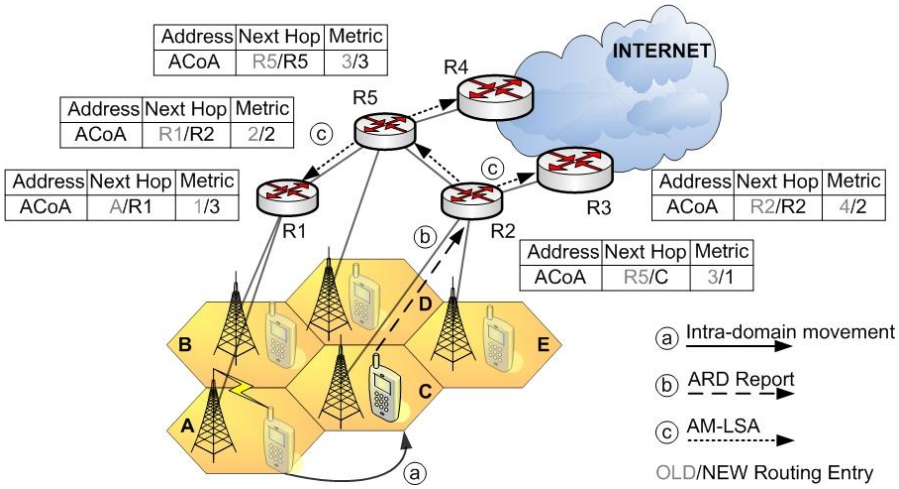


Figure 5: Details of AOSPFv3 operation in ABMF

2.1.2.3 Summary of ABMF

Several proposals are presented in the literature to deal with the performance problems of Mobile IP in micromobility scenarios. In [47] a comprehensive study is given on the performance of seven IP micromobility protocols and a performance analysis framework is presented consisting of five key performance indicators: handoff performance, passive connectivity and paging, intra-network traffic, scalability and robustness. This section summarizes the pros and cons of ABMF based on the analysis of [47].

ABMF is independent of the radio layer, and there is no need of implicit movement detection in the uplink direction: the mobile terminal may continuously transmit packages during handover. If the terminal enters a new Internet point of attachment (PoA), it still uses its old anycast address, which is valid in the whole logical subnetwork, meaning that there is no need for the time consuming address acquiring process during intra-domain handover events. In my proposal, the communication disruption in the uplink direction is only limited to the time requirements of Layer 2 procedures.

Downlink routing entries are automatically adjusted by a trigger of the first data packet arriving at the new PoA. When receiving a data packet from a mobile node, the access router changes the metric for the given anycast address in its routing table. This event initiates an ARI/AM-LSA message at the lowest level access routers, and the routing information spreads in the micromobility domain. Right after this update reaches the crossover router, the downlink data packets will be routed appropriately towards the new location of the mobile terminal.

Majority of existing micromobility protocols do not provide efficient support for intra-domain traffic. E.g., in case of CIP (Cellular IP) [11], all intra network traffic is routed through the CIP domain Gateway. As ABMF relies on IPv6 routing to drive packets towards

their destination, the intra-domain communication will be optimal, regardless of the destinations. The re-use of built-in IPv6 routing results in low number of involved stations during the handover and low handover latency. The low number of explicit management messages also decreases signaling overhead over the radio interface: the management load is rather shifted to the wired network segments with more resources available.

The most important advantage of my ABMF proposal is that there is no need to employ new protocol like in case of CIP [11] or HAWAII [48], but only the built-in capabilities of IPv6 and MIPv6 are used. Although the standardized operation of an IPv6 anycast routing protocol is currently work-in-progress and numerous anycast related research projects are still ongoing, it is very likely, that my solution will require only minimal modifications to the network if deployed.

The inside-domain movement of the mobile host is completely transparent in the direction of the Home Agent and the Correspondent Node when ABMF is applied, this could be useful when security and intractability is taken into consideration during network planning.

In order to reach an idle mobile node in mobile networks, some proposals (like [11] or [48]) include paging support. In ABMF, the default scheme to locate a mobile terminal is performed by flooding the network with ARD query messages. As it was pointed out, more sophisticated paging solutions can be easily implemented in ABMF to track and locate mobile terminals.

In ABMF, the anycast address as local identifier of mobile nodes has the ability to support either soft state handovers (i.e., when mobile nodes are connected to both the new and the old PoAs simultaneously) or hard handovers (i.e., when mobile nodes are connected to only one PoA at once), if the underlying IPv6 anycast infrastructure supports multi recipient anycast routing (bi-casting or n-casting).

Majority of existing micromobility protocols rely on hierarchical networking structure to reduce routing update latency. This fact results in vulnerability to failures and also to increasing overhead at higher hierarchy levels. ABMF is not sensitive for node or link failures, as it does not contain any centralized nodes or databases (like the CIP domain Gateway in [11], GMA in RegReg6 [67], or the MAP in HMIPv6 [9]): the routing information is distributed among the network nodes inside the micromobility domain, providing a highly decentralized scheme, in correspondence to the philosophy of IP communication.

ABMF does not introduce extra signaling load on the wireless interface. However, on the wired part ARI messages are intensively exchanged upon handover events. The signaling load in the wired network is caused by the routing information updates, which is proportional to the number and the handover frequency of mobile terminals in the domain, and similar to existing micromobility solutions following the hop-by-hop routing approach.

2.1.3 Simulated Annealing Based Anycast Subnet Forming

2.1.3.1 Introduction to micromobility domain planning

It is usually hard to design the size of a micromobility area (i.e., locally administrated domain). Several important questions arise: how to group wireless points of attachments with their relevant coverage into micromobility domains, what kind of principles must be used to configure the hierarchical levels if they are available, and in which hierarchical level is advisable to implement special functions (e.g., anchors or gateways). The traffic load and mobility of nodes may vary, therefore a fixed structure lacks of flexibility: design schemes are needed to comprise these network dynamics and to provide optimal or near-optimal solutions.

An obvious algorithm is to group those access nodes and their coverage areas (i.e., cells) into one domain, which has a high rate of handovers among each other. In that way the number of global location updates (Binding Updates/registration messages) can be significantly decreased. But joining too much access nodes into one domain would degrade the overall performance since it will generate a high traffic load on anchor/gateway nodes, and result in higher cost of packet delivery and paging. Contrarily a small number of cells/PoAs inside a micromobility domain will lead to a huge amount of location updates to the home network but will alleviate paging costs.

Based on these assumptions, He Xiaoning et al. [24] proposed a dynamic micromobility domain construction scheme which is able to dynamically compose each micromobility domain according to the aggregated traffic information of the network.

The related questions are very similar to the Location Area (LA) planning problem (where cells must be grouped into location areas in an optimal way [68], [69], as in micromobility domain planning we also need to search for a trade-off compromise between the location update and the packet delivery cost.

One of the most known LA planning schemes is the solution called Traffic-Based Static Location Area Design – TB-LAD [28] that groups cell pairs with higher inter-cell mobile traffic into the same LA. In this algorithm a list of neighbors is created for each cell, and the neighbor with the highest inter-cell traffic will be selected from the list and included in the same LA with this cell. In the next step the algorithm finds neighbors with the highest traffic from the neighbor lists of the cells that are included for the current LA and includes them into the current LA. This is terminated, when there are no more neighbors that can be included or the maximum number of cells is reached for the current LA. After this loop the algorithm starts the forming of the next LA in the same way.

However, in case of the Location Area Forming Algorithm – LAFA [70], LAs are not formed one after the other, but simultaneously, always including the actual cell-pair to an already existing LA or creating a new one, enabling to build the LA structure in a distributed way.

Based on the experiments of LAFA, the duet of the Greedy LA Forming Algorithm (GREAL) and the Simulated Annealing Based Location Area Forming Algorithm (SABLAF) was proposed in [25]. In this scheme GREAL is adopted to form a basic partition of cells into LAs in a greedy way without any additional assumptions for cell contraction, and then SABLAF is applied for getting the final partition. In [71] authors also propose a similar simulated annealing based LA planning method giving a heuristic and near-optimal solution for LA planning in tolerable run-times.

There is also a specific Location Area planning algorithm for GEO Mobile Satellite Systems: by the way of extensive comparison of the cost of location management using different types of location area designs, an appropriate scheme was separated by the authors satisfying the special requirements of GEO satellite systems [72].

There are also Location Area and micromobility domain planning algorithms which are able to handle network structures with hierarchical levels [26] [J3] for assignment of an optimal tree structure to a given source of access router handover rates.

2.1.3.2 Algorithm proposal for Simulated Annealing Based Anycast Subnet forming

However work-in-progress / under standardization IPv6 anycast routing protocols can be re-used for ABMF purposes, a serious concern when introducing ABMF (and any hop-by-hop micromobility solution) still exists: since mobile nodes must be maintained as separate routing entries, the size of routing tables in the routing domain can easily explode. In order to control the size of the routing domain, keep the scalability and help the design and formation

of micromobility domains in ABMF, I have proposed a special subnet optimization algorithm also handling the tradeoff between the paging cost and the registration cost.

Thesis I.2 [C9], [C12], [J3] *I have developed a two-phase anycast subnet forming algorithm where firstly a greedy grouping is adopted to form a basic partition of wireless attachment points into anycast subnets (ASs), and then simulated annealing is applied to provide the final partitioning. I have shown that the proposed two-phase Simulated Annealing Based Anycast Subnet forming algorithm (SABAS), which is an improvement of the SABLAF scheme, reduces the registration cost by an average 35% compared to the reference forming scheme.*

In ABMF, at each AS boundary crossing, the mobile nodes register their new locations through signalling messages of MIPv6 in order to update the location management database of the Home Agent. In this way the system is able to maintain the current location of each user, but this will produce a registration cost in the network. Therefore the question arises, what size the AS should be for reducing the cost of paging, maintaining routing tables (inter-domain handovers) and registration signalling (intra-domain handovers).

On the one hand if we join more and more wireless points of attachment with their relevant territory (e.g., cells in cellular networks, or Internet Points of Access – PoAs with a certain coverage in IP mobility terminology) into one anycast subnet, then the number of subnet handovers (inter-domain movements) will be smaller, so the number of MIPv6 binding update messages [6] sent to the upper levels will decrease. But in case of big number of PoAs or belonging to a subnet, more possible mobile nodes can join into one micromobility domain (increasing the possibility of routing table explosion), and an incoming call will cause lot of paging messages. On the other hand if we decrease the number of PoAs, then we do not need to send so much paging messages (hereby we will load less links and the processing time will decrease, too) and the scalability problem can be solved as well, but then the number of subnet changes will increase. Therefore the overall problem in subnet planning for ABMF comes from the tradeoff between the paging cost and the registration cost, also considering the scalability issues.

I qualified the paging cost and maximal routing table size as a constraint: therefore the registration cost was left alone in the objective function. Hence I defined and formulated a problem in which the final goal is the determination of optimum number of wireless Internet points of attachment per an anycast subnet for which the registration cost is minimal, with the limitations of the paging cost and the routing table sizes as an inequality constraint function.

This problem is similar to the well-known Location Area planning problem [27], [28], therefore I have applied the widely used fluid model [73] for calculations about the movement of MNs among the ASs, relied on the results of [74], [75] for the definition of the MIPv6 registration cost and the paging cost, and used the equation of [45] for the calculations of $N_{\max}$ (the maximum possible number of PoAs in the AS) as a main input for my AS forming algorithm.

2.1.3.3 The MIPv6 registration cost

By employing ABMF the PoA coverage boundary crossing inside the anycast subnet (AS) will be hidden from the upper levels, meaning that administrative messages for registering the new location of a mobile at the macromobility management protocol (i.e., MIPv6 Home Agent) will not be sent during intra-AS handover events. In order to make calculations about the movement of MNs among the different ASs, I have selected the fluid flow model [73]. The fluid flow model describes the aggregate mobility of the mobile nodes in a set of PoAs (e.g., an AS) as a flow of liquid. According to this model the MNs are

moving with an average speed v, and their direction of movement is uniformly distributed in the area. Therefore the rate of outflow from that area is described by [73] as

$$R_{out} = \frac{v \cdot \rho \cdot P}{\pi} \tag{1}$$

where v is the average speed of the MNs, ρ is the density of MNs in the area and P is the perimeter of the area. In this model it is very simple to define and analyze the MIPv6 macromobility registration cost function. The density of the MNs in an AS:

$$\rho = \frac{K}{N_k \cdot S} \tag{2}$$

where K is the number of MNs in the k th AS, N_k is the number of PoAs in the k th AS, and S is the area of a PoA. Every time when a MN crosses a coverage boundary of a PoA which is an AS boundary also, a macromobility registration process is started and a Binding Update message is sent to the Home Agent [6]. I consider here the intra-AS boundary crossing cost negligible, because intra-domain handover cost is not considered in the macromobility operation. Therefore I only need to determine the number of PoAs located on the boundary of the k th AS, like a subset of N_k, and the proportion of the PoAs coverage perimeter which contributes to the k th AS perimeter, similarly to [73]. Using this perimeter of the k th AS:

$$P_k = N_p \cdot \nu_p(N_k) \tag{3}$$

where N_p is the number of boundary PoAs and ν_p is the proportion of the PoA coverage perimeter in the AS perimeter in the function of N_k. The number of the boundary PoAs can be approximated as it has been done in [74]:

$$N_p = \kappa \cdot \sqrt{N_k} \tag{4}$$

The proportion of the PoA coverage perimeter which will be the part of the AS perimeter as well can be defined with an empirical relation [75]:

$$\nu_p(N_k) \approx \nu \cdot \left(a + b \cdot {N_k}^{\eta - 1}\right) \tag{5}$$

where ν is the perimeter of a PoA coverage and $a = 0.3333$, $b = 0.309$, $\eta = 0.574965$. Substituting N_p and $\nu_p(N_k)$ in (3), the expression for the perimeter of the k th anycast subnet becomes:

$$P_k = \kappa \cdot \sqrt{N_k} \cdot \nu \cdot \left(a + b \cdot {N_k}^{\eta - 1}\right) \tag{6}$$

Hence the number of crossing the k th AS boundary can be calculated by substituting the values of ρ and P_k in the outflow rate of the fluid flow model:

$$q_k = \left(\frac{\nu \cdot \frac{K}{N_k \cdot S} \cdot \kappa \cdot \sqrt{N_k} \cdot \nu \cdot \left(0.333 + 0.309 \cdot {N_k}^{-0.425}\right)}{\pi} \right) \tag{7}$$

As a MIPv6 registration process is started when the MN crosses a PoA coverage boundary which is boundary of an AS too, the total MIPv6 registration cost will be:

$$C_{\mathrm{Re}g_k} = C_{BU} \cdot q_k \tag{8}$$

$$C_{\mathrm{Re}g_k} = C_{BU \cdot} \cdot \nu \cdot K \cdot \kappa \cdot \nu \cdot \left(\frac{0.333 \cdot {N_k}^{-0.5} + 0.309 \cdot {N_k}^{-0.925}}{\pi \cdot S} \right) \tag{9}$$

where C_{BU} is the cost required for transmitting a MIPv6 Binding Update message.

2.1.3.4 The paging cost considering the routing scalability issues

To have a feasible network, the paging capacities should not be exceeded, therefore we need a paging constraint per an AS. The limited network capabilities of finding the exact location of an MN in case of incoming session will cause a limit on the peak session arrival rate, therefore I defined an upper paging cost constraint for every AS. Another constraint should be defined, the maximum number of MNs in one AS ($K_{\max}$), considering the challenges of the non-aggregatable anycast routing entries in a given anycast subnet. The paging cost for the k th AS should not exceed the paging cost constraint (the paging cost for the k th AS will be the sum of C_{P_i} over the N_k PoAs):

$$C_{P_k} = \sum_{i=1}^{N_k} C_{P_i} = \sum_{i=1}^{N_k} B_P \cdot N_k \cdot K \cdot \lambda_i < C_k \tag{10}$$

$$C_{P_k} = B_P \cdot N_k \cdot K \cdot \sum_{i=1}^{N_k} \lambda_i < C_k \tag{11}$$

where B_P is the cost of transmitting a paging message and λ is the number of incoming sessions of a MN. Here I assume that MNs have the same average number of incoming sessions in all the PoAs in the k th AS ($\lambda_i = \lambda$), so the paging cost will be:

$$C_{P_k} = B_P \cdot {N_k}^2 \cdot K \cdot \lambda < C_k . \tag{12}$$

2.1.3.5 Optimization of the MIPv6 registration cost

The goal is to determine the optimum number of PoAs per an AS for which the registration cost is minimum, with the paging cost and routing tables' limitation as an inequality constraint function. Assuming that the session arrivals (λ) follow a Poisson process and the function of the registration cost (9) is a monotonically decreasing function, the paging constraint can be written as:

$$P(C_{P_k} < C_k) < 1 - e^{-\gamma} \tag{13}$$

where $\gamma = (10,100)$, depending on the accuracy of the paging constraint. The monotonically decreasing attribute of the registration cost function means, that we need to find the highest value of the N_k for which the (13) will be still satisfied. Substituting the expression of the paging cost in (13):

$$P(B_P \cdot {N_k}^2 \cdot K \cdot \lambda < C_k) < 1 - e^{-\gamma} \tag{14}$$

If we still assume that the λ probability variable follows a Poisson process, then the maximum value of N_k can be calculated ($N_{\max}$):

$$P(\lambda < \frac{C_k}{B_P \cdot {N_k}^2 \cdot K}) = 1 - e^{-\gamma} \tag{15}$$

Substituting the calculated value of N_k in (9) provides the minimum of the registration cost. I will use this calculated N_k as an input for my anycast subnet forming algorithm.

2.1.3.6 The simulated annealing based AS forming algorithm

My goal is to develop an anycast subnet (AS) planning algorithm, which considers the paging constraint, the scalability issues of IPv6 anycasting, and takes the available mobility pattern and PoA coverage perimeter information as input, and finds an optimal or near optimal AS structure for which the registration cost will be minimum.

The registration cost is proportional to the number of handover events between different ASs (q), therefore the registration cost can be minimized by designing the ASs such that the PoAs belonging to one AS have the lowest boundary crossing rates among each other. I extended an existing realistic mobile environment simulator [45], [J3], which generates this boundary crossing database, a handover rate for each PoA pair, defined on the border of these PoAs. The incoming session statistic can be also generated for every PoA; therefore the paging cost can be calculated in the same time for every AS.

My proposed SABAS algorithm starts with a greedy solution, which will provide the basic AS partition as an input to the simulated annealing method. The algorithm chooses the PoA pair with the biggest handover rate in the given structure of wireless Internet points of attachment ($q_{\max}$) and includes the two PoAs into the AS_1 set of PoAs. In the next step, SABAS searches for the second biggest handover rate among the PoA pairs for which is true, that one of them belongs to the AS_1 set of PoAs. The algorithm checks whether the inequality

$$N_k < N_{\max} \tag{16}$$

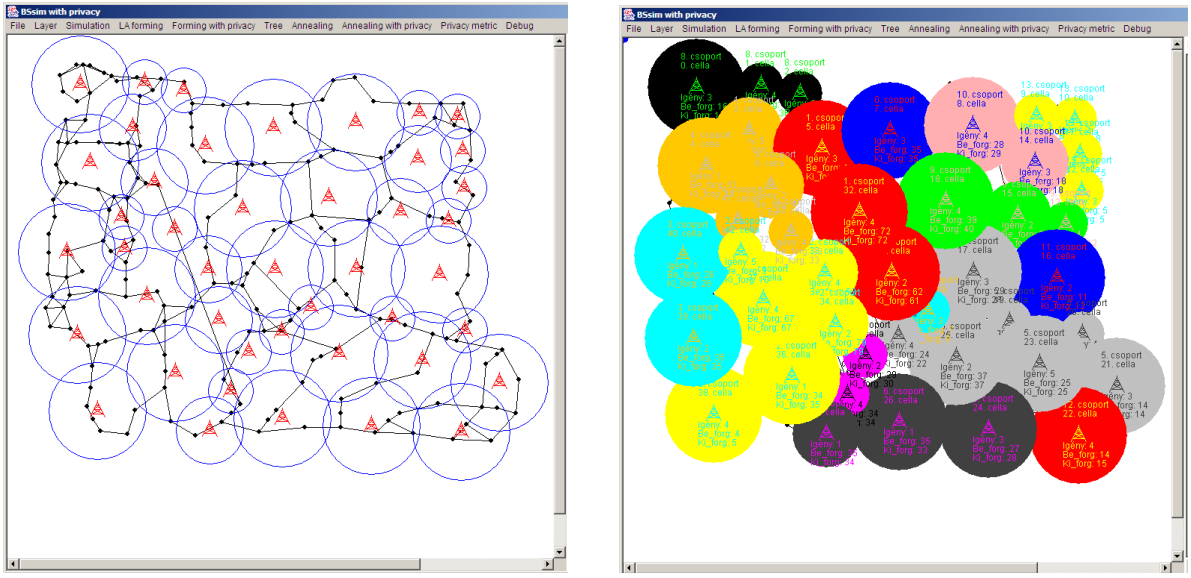
is satisfied, where $N_{\max}$ is the maximized value of N_k, namely the maximum number of PoAs in an AS which provides the minimum of the registration cost. If the inequality is satisfied, the PoA can be included into AS_1 set of PoAs. If the inequality is not satisfied, this PoA cannot be included into this set, not fulfilling the paging cost constraint. In this way SABAS joins the PoAs which are in the same dominant moving directions, therefore the number of handovers among ASs can be decreased (highways, footpaths, etc.). After the processing of all PoA pairs in the above sequential way, there will be PoAs that are not group of any set of PoAs. These PoAs will form another AS, which is not the best solution, but this will be only a basic AS partition which will serve as an input to the simulated annealing [76], [77] based AS forming scheme. The simulated annealing procedure starts with this basic partition, s_0. A neighbour to this solution s_1 is then generated as the next solution by simulated annealing, and the change in the registration cost $\Delta C_{\mathrm{Re}g}(s_0, s_1)$ is evaluated. The acceptance function is $e^{\left(-\frac{\Delta C_{\mathrm{Re}g}}{T}\right)}$, while the stopping rule is the maximal iteration step number or maximum number of steps when the $\Delta C_{\mathrm{Re}g}$ do not changes. I have defined another constraint, the maximum number of MNs in one AS ($K_{\max}$), considering the scalability challenges of the non-aggregatable anycast routing entries in a given anycast subnet. Therefore when the number of the routing entries reaches the $K_{\max}$ value in one AS (one routing entry for every MN), the value of the $N_{\max}$ need to be decreased, hence the ASs will consist of less number of PoAs in average, so the number of entries will be smaller in an AS proportionally. This decreasing should be continued until the number of routing entries goes under the $K_{\max}$ constraint.

2.1.3.7 Simulation framework and results

A Java-based realistic mobile environment simulator capable of providing rural and urban mobile environments [45], [J3] was extended by me in order to generate the input metrics of the algorithm. The simulator serves a two-fold purpose.

On the one hand it generates a realistic PoA coverage boundary crossing and incoming call (i.e., initiation of IP session) database in a mobile system given by the user with PoA, mobile node and movement path placing within the GUI. It also calculates both the handover rate for each PoA pair, defined on the border of these PoA coverages. The incoming session statistic can be also generated for every PoA; therefore the paging cost and the registration cost can be calculated in the same time for every domain.

On the other hand the simulator uses the above produced data as an input for the widest scale of location area and domain planning algorithms, and forms LAs and micromobility domains by running the implemented mathematical functions.



a) Example for initial cell/road structure in the GUI b) Example planned micromobility domain structure

Figure 6: The simulation software in use

The simulation can be executed on an arbitrary and customizable road grid covered by cells of various access technologies (e.g., Wi-Fi, GSM, UMTS) as shown in Fig. 6/a . Mobile nodes (MN) can be placed into this highly customizable environment by firstly specifying MNs' velocities, and setting the incoming session arrival parameter (IP session intensity).

This way different types of mobility environments can be designed (rural environment with highways or a densely populated urban environment with roads and carriageways, etc.,), together with the grids of cells configured and adapted to these environments. The applied mobility model here for MNs is the following. The different mobile terminals will move on the defined road grid by time-to-time choosing a random destination point on the road, similarly as in real life. Since typical mobile users are on the move aiming to manage a specific duty or reach a particular destination (e.g., heading to a hotel, a workplace, a hospital, etc.,) and they usually want to arrive in the shortest possible time, therefore the Dijkstra algorithm is used in the simulation framework in order to find the shortest path for mobile hosts towards their selected destination. The average speed of MN movements is defined by the velocity parameter of each mobile node.

For every mobile node an incoming session arrival parameter is defined and when a session initiation packet hits the node, the simulator designates it to the cell/PoA coverage where the node is in that moment. When a mobile host changes a cell or Poa, the simulator registers that a handover (i.e., boundary crossing) happened between the respective coverage-pair. When a simulation run ends, the simulator sums the cell/PoA boundary crossings and incoming session initiation distribution for every access network in the simulated topology. The results (road structure, cell structure, call numbers and cell matrix, mobile data) can be saved and opened to easily provide inputs for the Java implementation of the examined algorithms. An example domain structure gathered at the end of the whole simulation process is depicted in Fig. 6/b.

My goal with this mobility simulator was to evaluate my SABAS micromobility domain planning algorithm. I have compared SABAS with a manual AS grouping solution where the partitions are made intuitively (this reference manual solution should be considered as a

planed partition, but likely not the optimal one). I have examined how the registration cost changes by increasing the maximum number of PoAs in one AS.

Two environments were prepared in the simulator: the rural is rarely populated, but on the belonging highways a big number of mobile terminals are moving with high speeds, while the urban scenario is densely populated, with mobile terminals moving with smaller velocities. In the rural case the average PoA coverage size is larger then in the urban environment, accordingly there is a smaller number of PoAs. The rural environment consisted of 42 PoAs, while in the urban network there was 79 PoAs. I have executed the simulation on these two network scenarios and used the output of the simulation (boundary crossing and incoming session database) as the input parameter for SABAS. I have analyzed how the registration cost changes by increasing the maximum number of PoAs in one anycast subnet.

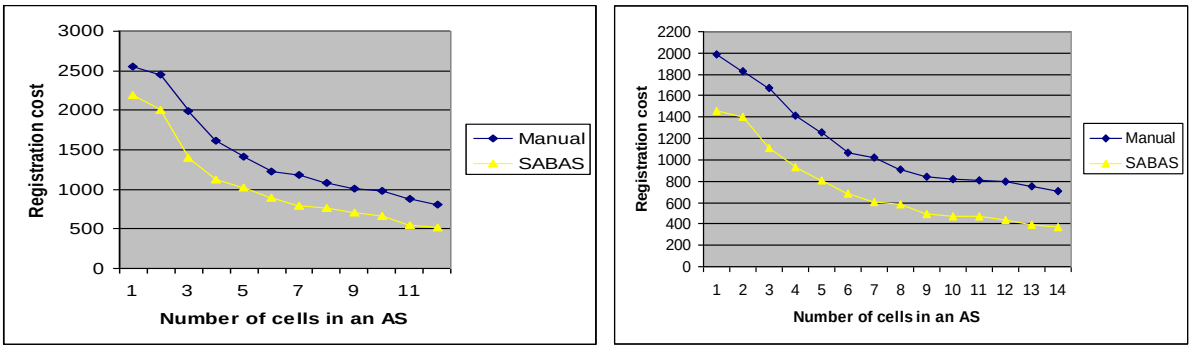


Figure 7: The registration cost in rural (left) and urban (right) environments

Fig. 7 left shows the registration cost in rural environment, where the *x* axis represents the $N_{\max}$, used in the (16). As it is depicted, SABAS finds a much better solution for every value of $N_{\max}$ than the manual solution. For the calculated value of $N_{\max} = 12$, the registration cost reaches the global minimum value using the proposed technique.

Fig. 7 right shows the registration cost in the urban scenario, where we have more PoAs, but the size of the cells or PoA coverages is smaller. In urban environment also, the SABAS is outperforming the manual solution significantly. Using again (16), $N_{\max} = 14$ (higher value than for the rural environment, because of the PoA coverage sizes and numbers), for that value the SABAS algorithm gives again the minimum of the registration cost

Summarizing the results depicted in Fig. 7 I can state that SABAS finds a much better solution for every value of $N_{\max}$ both in rural and urban environments, and decreases the average registration cost by an average 35% compared to the reference algorithm. This significant signalling load reduction in the domain structure can help us improving performance of QoS-sensitive applications while also taking care of scalability issues of my anycast based micromobility framework.

2.2. HIP-based Micromobility Management

Despite the fact that the today's Internet architecture is quite robust and supports innovation and rejuvenation with the proliferation of IPv6 technologies, the basics of TCP/IP have not changed since the first RFCs published in the early 1980's. The Internet Protocol was not designed with any kind of mobility in mind: the inseparable bond between the locator (Loc) and identifier (ID) functions of IP addresses makes it complicated, inconvenient or in some cases even impossible to design efficient, scalable and secure mobility and multihoming

solutions. To get down to the roots of this problem by separating the dual role of IP addresses and provide an extended TCP/IP stack for future mobile Internet, the Host Identity Protocol (HIP) [20], [21] was designed. The most important characteristic of HIP is that it uses cryptographic keys to identify hosts or other network entities while legacy IP addresses are acting as pure locators. In this architecture transport level connections are not bound to IP addresses, which are dynamically changeable in several cases, but to permanent identifiers, which remain the same for quite a long time. This property provides sophisticated and secure mobility/multihoming support [22], [23], [78] for standard macromobility scenarios, but further extension of the base protocol is needed for micromobility scenarios. The original idea of integrating micromobility with HIP was presented in [79] but their solution was not built on an effective and intact micromobility model as the focus was on the security issues, and the authors did not consider protocol details regarding the operation and the mobility support. This motivated me to develop a complete micromobility solution based on Host Identity Protocol, and by this way to highlight the emerging mobility applications of this promising ID/Loc separation protocol family.

2.2.1 HIP in a Nutshell

2.2.1.1 The Host Identity Layer

Current IP networks are based on two basic kinds of namespaces. On the one hand there are human readable domain names which can be resolved to IP addresses by Internet applications via Domain Name System (DNS) lookups. DNS provides fast queries but it is not designed for fast updates and quick retrieval of dynamic information. On the other hand there are IP addresses used in the network layer as locator for packet routing purposes and also used as identifier in upper layers to refer to the host or a particular communication session. The general concept of ID/Loc separation aims to eliminate the problems and limitations by splitting the two roles of IP addresses and such allowing network layer to change locators without interfering with upper layer procedures. This separation makes the routing infrastructure more scalable, and by introducing a mapping function between IDs and Locs a natural and effective support of mobility and multihoming can be provided.

The concept gains more and more popularity: several different approaches exist for ID/Loc separation and it also has recently been introduced in the standardization activities of the ITU Telecommunication Standardization Sector (ITU-T) for integration in future network architectures [18], [80]. The common in all the existing standards, drafts and recommendations is the separation of identifiers from locators and applying a dynamic mapping mechanism between them, making the duplicate role of IP addresses disappear. They either use distinct namespaces for both functions (i.e., ID and Loc) or provide an architecture where the nature of the split is operational.

Host Identity Protocol uses the first approach: IP addresses continue to act as pure locators, while the identification role is handled by a newly introduced, globally unique namespace (the Host Identity namespace), that is a special pool of identity representations called Host Identifiers (HIs). The elements of the Host Identity namespace are public keys of asymmetric key pairs (i.e., self-certifying cryptographic names) used to identify nodes and to integrate strong security features such as authentication, confidentiality, integrity and protection against certain kind of Denial-of-Service (DoS) and Man-in-the-Middle (MitM) attacks. Furthermore, based on the cryptographic HIs special certificates can be generated by the nodes for secure signaling [81] or even identity delegation [82], offering enormous resource savings, effective session mobility and other promising application possibilities in wireless and mobile environments. However, HIs are rarely used in actual HIP protocol

packets, instead hash representations called the Host Identity Tag (HIT – 128 bit long for global, IPv6-based communication) and Local Scope Identifier (LSI – 32 bit long for local usage and IPv4 compatibility) are applied. HIP related signaling information is conveyed within HIP headers having a form of a standard IPv6 extension header. Every HIP compatible node has at least one HI and implements the functions required to handle the new namespace and the relevant mechanisms. Therefore the scope of the protocol includes the modifications and new methods designed to integrate the concepts of HIP into the existing Internet architecture. As Fig. 8 shows, these functions and TCP/IP extensions form a new protocol layer, which resides between the transport and network layer in the TCP/IP reference model [21].

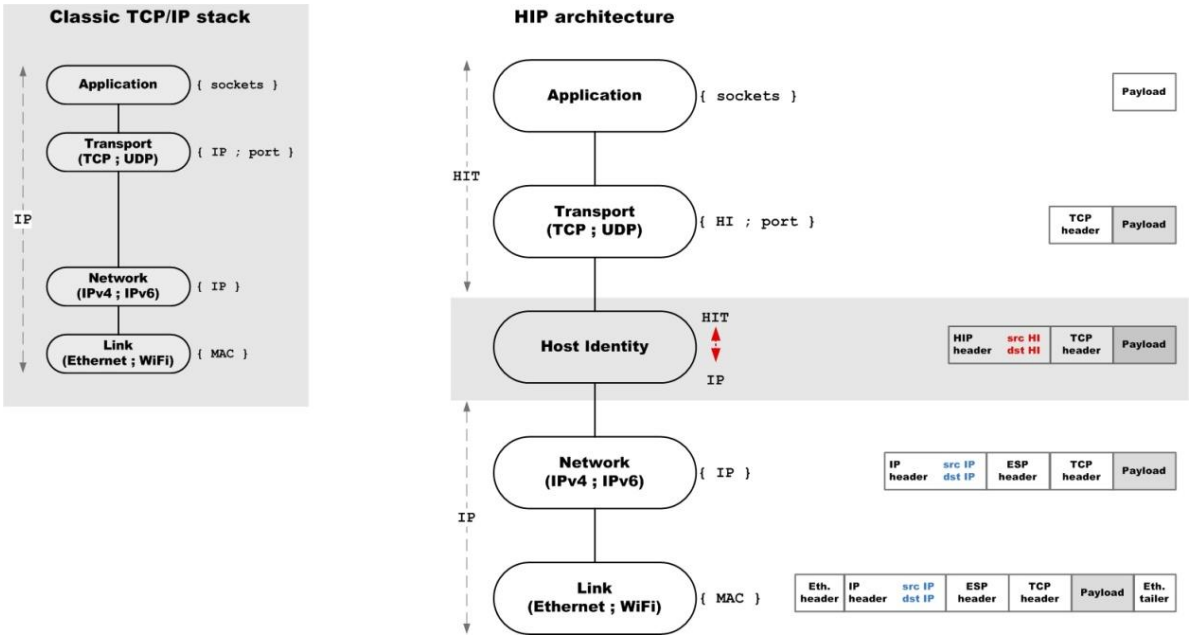


Figure 8: The Host Identity Layer

The basic function of HIP is to set up Host Identifier based connections between nodes and to map HIs to IP addresses and vice versa. A HIP association can be established between two nodes (i.e., an Initiator and a Responder) by a four way end-to-end security handshake called the Base Exchange (BEX) (see Fig. 9).

The BEX performs mutual authentication based on the peers' asymmetric keys and implements a Diffie-Hellman key exchange to create symmetric keys for later payload encryption. Additionally, a special puzzle-solution mechanism is applied to protect the responder against certain DoS attacks. As a result of a successful HIP Base Exchange an IPsec Security Association pair is created between the peers where SAs are bound to HIs instead of IP addresses [20]. After the BEX, payload data is passed between the peers using the Encapsulating Security Payload (ESP) through a special ESP tunnel. A new transport mode of ESP was designed especially for HIP [83]. This so called Bound-End-to-End-Tunnel (BEET) mode integrates the ESP tunnel mode with the low overhead transport mode. Using BEET mode the outer IP header of the ESP packet holds the IP addresses of the peers but the inner header is missing. Instead the Security Parameter Index (SPI) is used to identify the correspondent HIP association by reception at the destination. Thanks to the BEET mechanisms HIP related signaling information (i.e., HIP header with source and destination HITs, and HIP parameters) must be applied only to HIP control packets but not in case of data transfer messages.

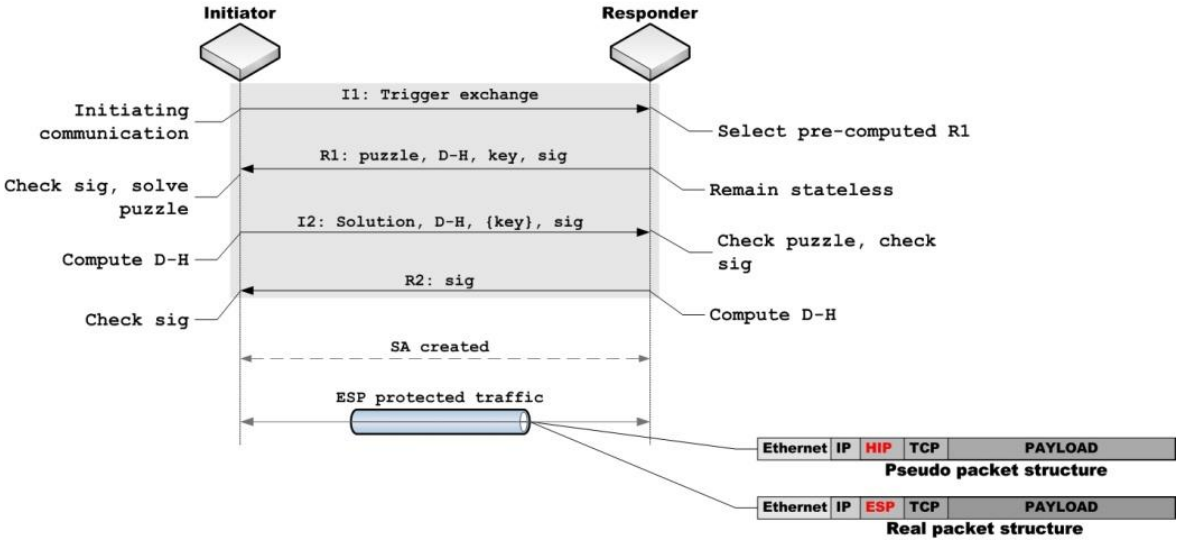


Figure 9: The HIP Base Exchange

During HIP operation IP addresses (i.e., locators) are intended to be used mostly for ”on-the-wire communication“ between peer hosts, while upper layer protocols and applications use HIs (or HITs) instead. This implies the need of some method to translate domain names to HIs. Using the existing infrastructure of DNS for this translation is quite straightforward. Therefore in [84] authors designed a new resource record for the DNS, and laid down how to use it with the Host Identity Protocol. This novel resource record allows a HIP node to store its Host Identity and other relevant information (e.g., HIT) in the DNS.

2.2.1.2 Basic HIP mobility and multihoming support

The base specification of HIP describes a secure locator update procedure, which I describe here in detail. The procedure is used to maintain the HIT-IP mappings between the communicating peers [22]. The mobile endpoint informs partners that its location has changed. Inherited from the key idea of HIP the update procedure does not affect higher layer connection. The procedure is transparent for all established connections of the transport or application layers. This property makes HIP an exciting ground to develop sophisticated mobility schemes or use it to handle more complicated and advanced mobility scenarios. The update sequence is illustrated on Fig. 10 (left). This is the most simple mobility scenario specified for HIP. There are two HIP capable nodes, which have established communication sessions. Note that their higher layer connections are bound to HITs instead of IP addresses. In case the IP address of the mobile node is changed, it will trigger a HIP update procedure by sending an UPDATE (U) message to its peer(s). This delivers the new location information (loc) and informs the peer if the mobile wants to update the security parameters (esp). If there is a need for refresh, the mobile also sends the updated parameters (D-H). The update procedure is proved to be protected against security attacks. On the one hand all the messages are digitally signed by the peers, authenticating the origin of the message and the message for any party using the HI of the sender. On the other hand there is a built in protection against distributed Denial-of-Service attacks. The second and the third message of the update procedure implements this. The peer node receiving the first UPDATE packet verifies the signature and answers with another UPDATE packet. This includes information to update the security parameters (esp and D-H) and a data block that contains a nonce (e_req). This must be echoed back by the mobile node in the third UPDATE packet (e_res). This simple echo request-response sequence verifies the new address of the mobile node.

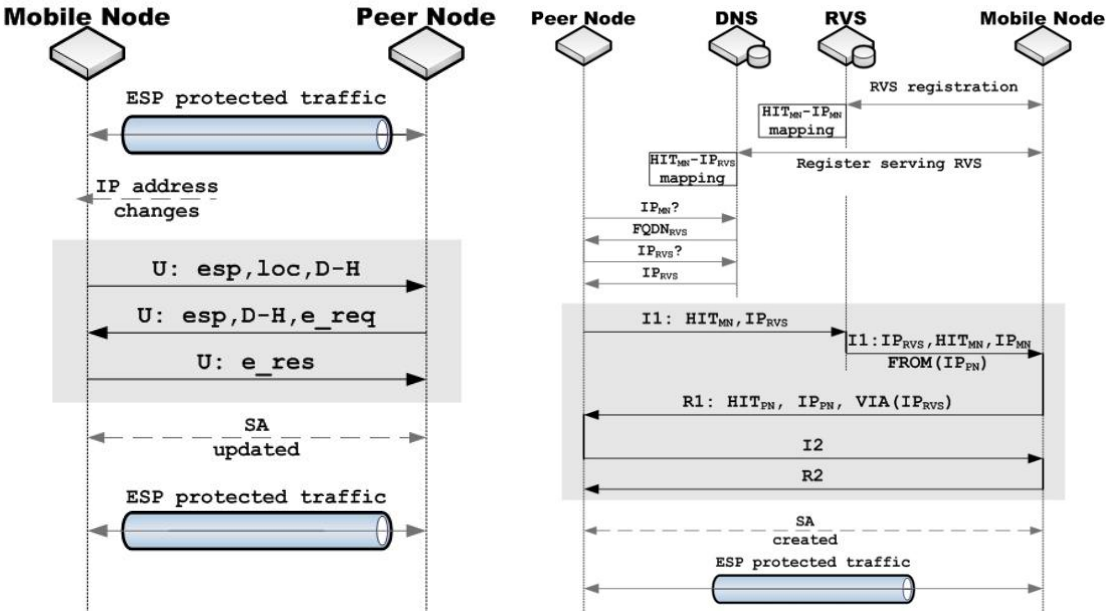


Figure 10: The HIP UPDATE procedure (left) and the HIP RVS mechanism (right)

A related functionality of HIP is host multihoming. In case of multihoming the HIP node owns more than one physical interfaces and/or global addresses. However the update procedure described above is used to update the primary locator of HIP nodes, a multihomed node can inform its peers about secondary locators it is reachable at. It is recommended to use different SAs for different interfaces and/or addresses. To do this, a multihomed HIP host creates a new inbound SA and a corresponding SPI. This is also managed by the update procedure. The first UPDATE packet should hold an ESP_INFO parameter having the NEW SPI field set to the newly created SPI value and setting the OLD SPI field set to zero. The packet also contains a LOCATOR parameter that indicates the new address-SPI mapping and the old one as well. Peers will use the primary locator as long as it is available and can switch to one of the secondary locators upon loss of connection.

The above introduced update procedure for mobility and multihoming can handle locator changes in case there are ongoing HIP sessions between the endpoints. However this is not a solution for initial reachability of mobile nodes and cannot cope with simultaneous mobility of endpoints. Initial reachability is the problem about how to provide a permanent anchor point for mobile nodes that makes it able to reach them no matter what their actual location is. Simultaneous end-host mobility covers scenarios where both endpoints are moving away from their location more or less parallel. Thanks to this the UPDATE messages cannot reach their destination. The messages are delivered to the old locations of the peers and the partners will lose each other and they have to restart their common session(s). There is an extension in HIP standards that introduces an anchor point called the Rendezvous Server (RVS) to solve the above cases [23]. Fig. 10 (right) shows the service the RVS provides for mobile HIP nodes to handle scenarios described above.

The RVS is known to every potential peer nodes by e.g., DNS queries. The RVS stores the actual HIT-IP mapping for registered mobile nodes. The Base Exchange is assisted by the RVS to enable connection establishment for the peers. Here we describe the sequence in detail. First the mobile node has to register itself at the RVS to use the offered service [78]. This creates an entry in the RVS database that holds the HIT-IP mapping for the mobile. The entry is updated time-to-time by the mobile if its IP address changes. After registration the mobile informs the DNS indicating its serving RVS. At this point any potential peer can initiate a HIP connection with the mobile. The peer performs DNS queries to get to know the serving RVS of the mobile it wants to reach. This is a two-stage query. First the peer asks the

IP of the mobile node indicating its domain name. The DNS returns the domain name of the RVS. In the second stage the peer asks the IP address of the RVS and the DNS returns its IP address. Now the peer can trigger the Base Exchange by sending the I1 message. This is delivered to the RVS. The anchor point knows the actual IP address of the mobile node and modifies the I1 message accordingly. The RVS also attaches an additional data field to the message that identifies the original sender of the message (FROM(IP_{PN})). The message is delivered than to the mobile, which continues the Base Exchange by sending the R1 message. This also contains an additional parameter, which verifies that the I1 message was forwarded by the RVS (VIA(IP_{RVS})). Finally the connection setup finishes in the regular way without the inclusion of the RVS. Note that the RVS is used only in the connection setup. Any other communication (signaling or data) between the peers is transferred in the direct path. The similar process must be followed when the endpoints are changing IP addresses parallel. In this case the HIP connection is broken and must be reestablished.

2.2.2 µHIP: Micromobility in the Host Identity Layer

Basic HIP mobility and multihoming mechanisms are only for macromobility support and further extension of the base protocol is needed for micromobility and other distributed mobility scenarios. The original idea of providing micromobility support in HIP was presented by authors of [79] where a secure micromobility management scheme based on Lamport one-way hash chains and secret splitting techniques was introduced for IP networks. However, this scheme is not built on an effective and intact micromobility model as it only focuses on the security issues and it also does not consider protocol details regarding the operation and the mobility support. Moreover, in their method MNs still need to update their location information at the RVS during the handover, therefore the scheme cannot fulfill the requirements of micromobility architecture: it is only a partial answer for the complex problem.

Thesis I.3. [C4], [C11], [C17], [C21], [B3], [J14], [J20] *I have developed a Host Identity Protocol based micromobility solution (µHIP) that makes HIP able to efficiently serve frequently moving mobile users while preserving all the advantages of the standard HIP protocol suite. I have also introduced a paging method fitting into the proposed µHIP architecture. I have shown by extensive simulations built on complex protocol models that my proposed µHIP scheme outperforms the standard HIP mobility management solution in micromobility environments by providing an average TCP performance gain of 20%, while introducing only a 9% decrease during the much less frequent macromobility scenarios.*

In order to distribute HIP anchor nodes (Rendezvous Servers – RVSs [23]) and control micromobility domains in the µHIP architecture I have introduced a novel HIP gateway entity called the Local Rendezvous Server (LRVS) which is responsible for managing HIP Mobile Nodes (MNs) in a given domain (Fig. 11). LRVS gateways provide HIP registration service for users in the domain, and also introduce an IP address mapping function which is used to attach the MNs to the µHIP access network by registering the local locators (IP_L) of MNs. IP_L is valid only in the given domain and the LRVS is responsible for mapping every IP_L to a globally routable address (i.e., global locator, IP_G). IP_G is used to register the MNs at their standard RVSs and to deliver packets outside the micromobility domain during further communication sessions.

2.2.2.1 Initiation mechanism

If a MN joins a new, locally managed network area, there is a need for an initialization mechanism in order to start the inner-domain life of the mobile node (see Fig. 12). After entering, the MN physically connects to one of the access routers (AR) of the domain. Right after detecting the newly established physical connection and getting a serviceable IP address (IP_L), the MN either may actively initiate a HIP service discovery procedure or passively wait for a service announcement in order to detect the LRVS service provided in the visited network area [85]. Irrespectively of the used mechanism the MN will eventually be informed about the HIT and the IP address of the LRVS responsible for the actual domain. Steps 1-3 in Fig. 12 (yellow arrows) show the case of passive discovery where the MN (due to its movement) sends an UPDATE packet to its RVS and current Correspondent Nodes (CNs). The LRVS – according to the procedures of passive discovery – intercepts the UDPATE packet, verifies the I1 source HIT and sends back a Service Announcement Packet (SAP) to the MN containing R1 data and information about the LRVS services (HIT and IP address of LRVS). After that the MN continues the service discovery by completing the registration to the LRVS with the final I2-R2 sequence. Till this point everything works almost the same way as it would be with a normal RVS. The main difference is that during this service discovery and registration procedure the LRVS not only opens a new entry in its database and registers the MN's HIT with its new, local IP address but maps it with a globally routable IP address (IP_G) as well.

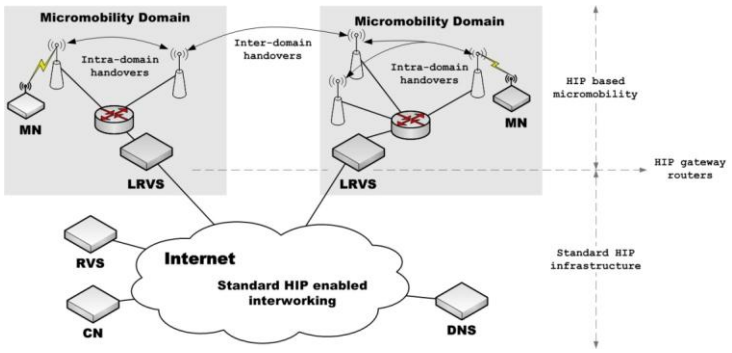


Figure 11: The proposed μHIP architecture

After the MN is registered at the LRVS, it needs to perform the update or registration at the RVS and its current CNs as well; to be reachable for the current and future communication partners (step 4, yellow arrow). Therefore the MN – strongly relying on the self-certifying cryptographic identifiers provided by HIP – delegates its signaling rights [81] to the LRVS at which it is registered. The appropriate certificates can be sent after the BEX between the MN and the LRVS, resulting that the LRVS will own the rights to signal on behalf of all mobile nodes in the current micromobility domain. In possession of this delegation the LRVS is able to securely register or update to the RVSs and CNs on behalf of the MNs with globally routable IP addresses assigned to them.

After this initiation procedure the MN is registered at the LRVS (with the HIT_{MN}-IP_L-IP_G triplet) and at the RVS (with the HIT_{MN}-IP_G pair) as well.

If a node (MN_2 in Fig. 12) that had performed the same initialization mechanism in a different domain wants to establish a HIP association with an also initialized MN_3, it sends the first packet (I1) of the Basic Exchange (step 1, blue arrow). In this packet the source IP address is the local IP address of the initiatior (i.e. MN_2), the destination IP address is the IP address of the MN_3's RVS (here we assume that the RVS of MN_2 and MN_3 are identical), the destination HIT is the HIT of MN_3. The I1 packet is intercepted by the LRVS of the initiator's domain (i.e. $LRVS_2$). This LRVS changes the source IP address of the packet to the globally

routable IP address of MN_2 (IP_{G2}) and sends the packet to the RVS (step 2, blue arrow). The RVS forwards the packet towards the registered (i.e., IP_{G3}) address of MN_3 (step 3). $LRVS_3$ knows the actual attach point of MN_3, so it forwards the packet by changing the destination IP address of the packet (IP_{G3}) to the MN_3's local address (IP_{L3}) (step 4, blue arrow). The BEX continues in the regular way, without the inclusion of the RVS, but with the address changing function of the two LRVSs (step 5, blue arrow). (Note that the CHECKSUM field of IP packets should be recomputed after every address change, similarly as in case of standard RVSs.)

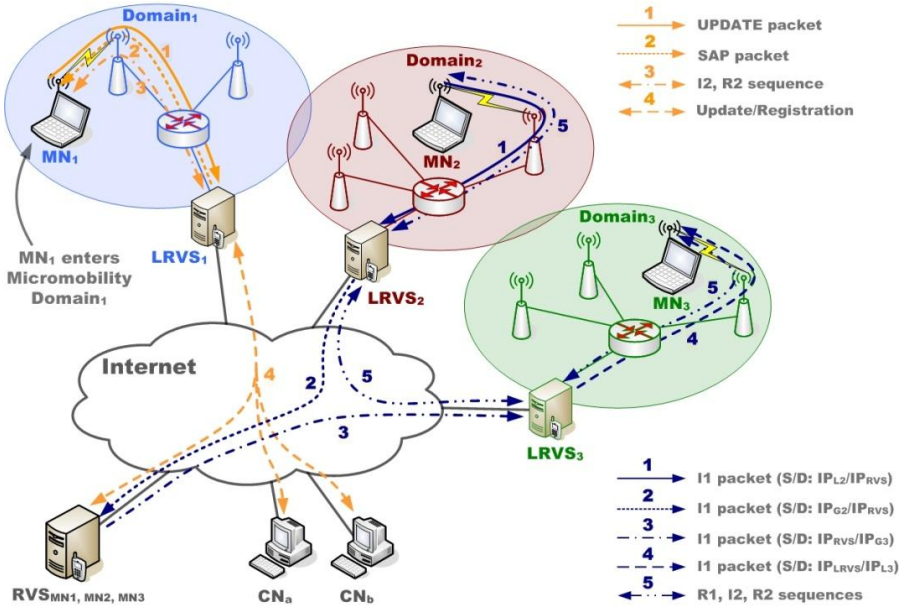


Figure 12: Initiation mechanism and connection establishment in the µHIP framework

After this message sequence there is an active HIP association available between the two nodes, and they can begin sending data packets to each other. Data packets are forwarded by the LRVSs to the actual local IP addresses of the MNs in the same way as they did during the BEX. It is crucial to observe that due to the HIP based signaling delegation all the above functions of the LRVS system are to be considered secure.

2.2.2.2 Intra-domain handover procedures

If a moving node which had performed the initialization mechanism described in the previous section, moves to another possible point of attachment of the same domain, the MN will receive a new IP_L from a servicing AR belonging to the same LRVS (see MN_1 and yellow arrows in Fig. 13). In this case the MN – realizing the change of its IP address – simply updates its registration (and if needed its delegation certificate as well) with its new local IP address at the LRVS. The used update mechanism is the standardized way detailed in [78]. It is important to note that neither the CNs of the MN nor the RVS has to be informed about the movement as it is locally handled. The address changes within a domain are managed by the LRVS system responsible for that particular domain: the movements of the MN are completely hidden from the outside world in order to reduce the signaling overhead, packet loss and handover latency in a significant degree.

2.2.2.3 Inter-domain handover procedures

MN_2 and blue arrows in Fig. 13 demonstrate a possible inter-domain handover scenario. In such cases, the MN moves between local administrative domains (i.e., $Domain_2$ and $Domain_3$) thus invoking global mobility management procedures of µHIP. Arriving at the

new domain, the MN will receive its new IP_L, and will discover the service parameters of the new LRVS ($LRVS_3$). After the MN realized that it leaved the previously used administrative domain by entering a new one and learned the HIT and IP address of the new LRVS, it performs a registration mechanism. This works the same way, as we described in Section 2.2.2.1. Since MN changes its old LRVS, it has to update its RVS and all the correspondent nodes with ongoing communication. But the first thing to do is to update the old LRVS (i.e., $LRVS_2$) to make it able to forward packets sent to the MN's old globally routable IP address as long as the MN has not finished updating the RVS and all of its CNs (step 3, blue arrow). After the old LRVS is updated, its CNs and at last the RVS will also be updated (step 4, blue arrows). This update informs the RVS about the MN's new IP_G, which was given by the new LRVS (i.e. $LRVS_3$). When all of the required updates are finished, the registration association at the old LRVS will be removed (or automatically timeouted).

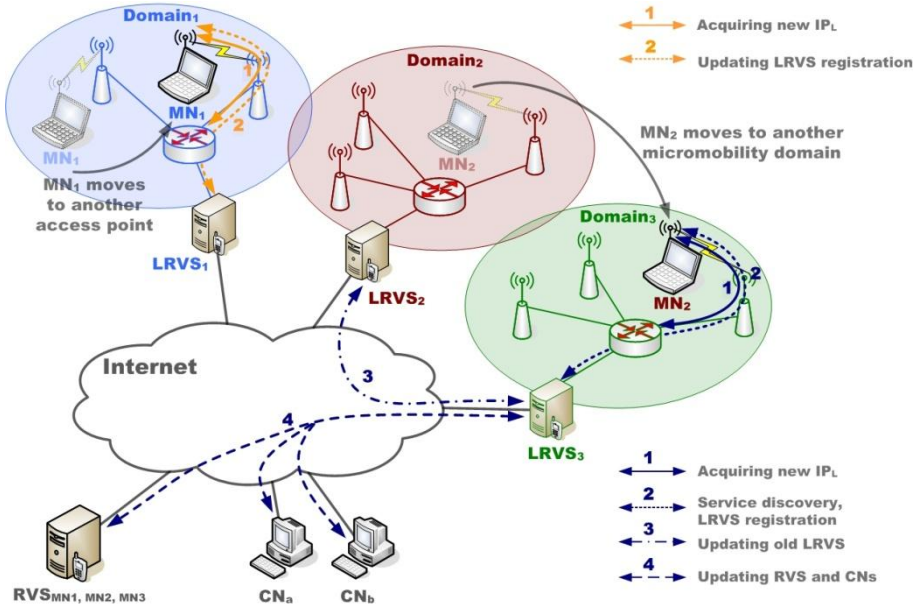


Figure 13: Intra-, and inter-domain handover procedures

2.2.2.4 Paging

In mobility supporting IP based networks the exact topological location of every mobile node must be known for appropriate packet delivery. Therefore a serious trade-off has to be considered namely how tightly the network should track the actual location of a mobile node (i.e. how frequently should a MN send location updates) versus the required resources to locate a particular mobile node whose current position is not accurately known. In order to make possible to deal with this trade-off in our proactive micromobility supporting framework, a HIT specific multicasting based paging extension is to be introduced in the system.

The basic idea of my proposed paging scheme is shown in Fig. 14 via an example μHIP network configuration. The case of MN_1 shows the extended initialization mechanism with paging support, while the case of MN_2 introduces the procedures when a correspondent node from the outside network initiates a transmission towards an already initialized but actually inactive MN (MN_2) residing inside the domain. During the extended initialization the MN registers itself into the currently entered Paging Area (PA) as well. If there are no ongoing communication sessions on a registered MN, the mobile node only needs to update its LRVS when it migrates into another Paging Area. Therefore when an incoming session is detected in the LRVS (i.e., CN's packets are reaching the designated MN's LRVS), and if the LRVS

system doesn't know the exact location of the destined MN (i.e., the MN_2 is in standby mode for a long time and its registration information is outdated, only its Paging Area is known), then it triggers the paging mechanism: appropriate paging requests will be sent by the LRVS system towards all access routers within the designated MNs suggested location area determined using a Paging Registration Database. Transporting the paging requests is done by a simplified HIT specific multicasting method based on [86], that carries the requests towards the MN through the corresponding subset of routers. The MN will be eventually reached thus forcing it to perform a registration with the LRVS which results in successful connection build up with the initiator CN.

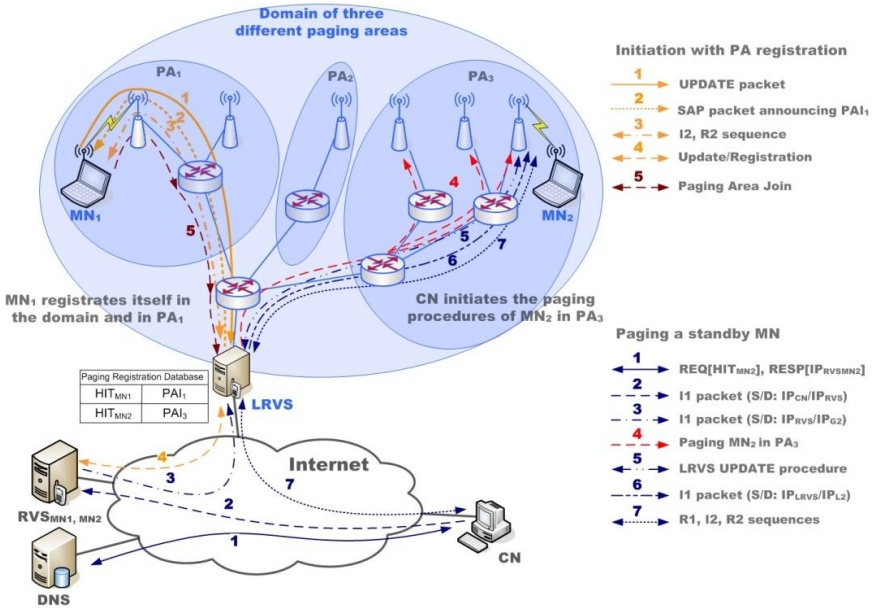


Figure 14: Mechanisms of paging in the µHIP framework

To support this operation a Paging Registration Database (PRD) located in the LRVS system is introduced. Every record in the PRD contains a Host Identity Tag (HIT) – Paging Area Identifier (PAI) pair. In my proposal PAIs are multicast addresses uniquely assigned to every paging area. Paging Area updates (i.e., PRD mappings) have longer timeout compared to a normal LRVS registration lifetime implicating a longer interval between consecutive PA updates. Note that both a normal LRVS registration/reregistration and a PRD mapping/remapping sequence can be initiated by sending a standard HIP UDPATE packet with an included LOCATOR parameter, but the different requests should be distinguished with different REG_INFO content (e.g., using the Reg_Type field for differentiating between the two functions) [78]. Thus the Paging Registration Database can be maintained and kept updated similarly to the LRVS registration synchronization mechanisms introduced in the previous chapter.

The carrying of the paging messages is based on [86] where authors present a Host Identity Specific Multicast (HISM) model and a unified Version Independent Group Management Protocol capable of handling HITs in order to provide solution for access control, accounting, mobility, and IPv4/v6 transition relating problems of the conventional multicasting methods. Based on their model I introduce a simple method which perfectly suits for the requirements of carrying the paging messages in our HIP based micromobility framework.

This method assumes a Mobile Node, which after entering a µHIP domain and finishing its initialization duties is considered to be registered in the LRVS system. However, in order to make prepared the MN for the paging functions we should integrate some other procedures

into the initialization mechanisms. The scheme with the extended initialization is the following (Fig. 14, case of MN_1). During the LRVS Service Discovery sequence, the MN must be informed about its current paging area. The required information is the multicast group's IP address uniquely assigned to every single paging area, and can be included into the LRVS's SAP packet using the REG_INFO parameter (step 2, yellow arrow). After approving all the necessary information, the μHIP implementation registers the multicast address of the current paging area in the operation system's registry and prepares the Paging Area Join message. This message is a multicast group management Join message containing the HIT_{LRVS} as we are to create a source specific multicast tree based on HIT information, and the HIT_{MN} for possible supplemental authentication purposes according to [86]. The Join message arrives to the first multicast router which starts the multicast tree building procedures (step 5, yellow arrows). After reaching the multicast router of the LRVS system (or any other intermediate router which is already on a branch of the issued multicast tree) the paging area join procedure finishes and the MN is ready to be paged.

Note that our μHIP paging scheme doesn't require implementation of HIP layers in the routers even though the multicast tree is built based on HIT information. The only need is to implement the unified group management protocol, and using HITs in the routing table entries as tree states. All of these changes can be done by upgrading the router's multicast software.

2.2.3 Simulation environment and evaluation results

In order to evaluate my proposed HIP-based micromobility solution and to provide a highly configurable, extensible, and adequate model for HIP, μHIP and other related protocols, I have designed an IPv6-based Host Identity Protocol simulation framework called HIPSim++ [C17]. The model is built on the top of the 1.99.3 version of INET [87] which is an extension and TCP/IP model collection of the component based, modular OMNeT++ 4.2 discrete event simulation environment [46].

2.2.3.1 Details of the HIPSim++ simulation model

Considering that the 32bit LSIs are designed for local communication only (i.e., the benefits of the HIP scheme can't be exploited totally on IPv4), the designed HIP simulation framework uses the IPv6 networking stack of INET [87] such fulfilling the requirements of global HIP communication based on the 128bit HITs. Transparency of the novel HIP layer is another important requirement which should be guaranteed for practical reasons, but current HIP RFCs define only a few guidelines to support the transparent behavior of HIP in the current networking architecture. These guidelines were also followed during the development work of HIPSim++ [C17].

Despite the fact that HIP relies on the functions of IPSec, a full implementation of IPSec and relating algorithms is not part of my simulation model: HIPSim++ does not possess properly realized Diffie–Hellman mechanisms, RSA engine, cryptographic hash functions and puzzles because precise mapping of all the security algorithms is out of scope of my evaluation efforts. The main design goal of HIPSim++ was to accurately simulate core HIP instruments focusing on the advanced mobility and multihoming capabilities and wireless behaviour of the protocol and providing only skeleton implementation of the above mentioned mathematical apparatus.

The simplest scenario of introducing HIP into the ISO-OSI architecture is when applications continue to use IP addresses, and HITs (or LSIs) only appear in the newly introduced HIP layer. Besides the integration of the Host Identity Layer, no other modifications are to be applied in the current protocol stack if such a scenario is implemented.

However this is an easy way to introduce basic HIP functions, it also restricts HIP's general benefits of mobility and multihoming support. Therefore the implemented HIP layer in HIPSim++ registers HIT-IP bonds for every communication session, and when packets from the transport layer arrive, destination and source HITs are replaced by destination and source IP addresses. Higher layers know only about HITs and Port numbers: they are using HITs instead of IP addresses. By realizing this scenario, all the advantages and benefits of applying HIP can be exploited and also HIPSim++ can be easily used in the existing INET-based simulation models by implementing dedicated new modules for HIP specific functions.

The core of HIPSim++ is the HIP layer module named as *HIP module* which creates a daemon instance called *HIP SM* for every new HIP session. This daemon is responsible for all mechanisms of the HIP State Machine (HIP SM) described in [20], e.g., for handling HIP Base Exchange and HIP mobility functions. One such daemon instance cares of one SA, which will be identified by the local SPI. *HIP SM* daemons are registered by destination and source HITs (and SPIs) in the *HIP module*. HITs have to be provided by the applications (or rather the transport layer), therefore HIP-capable DNS extensions [84] are also integrated into HIPSim++. The HIP module is also responsible for managing changes occurring in the states and addresses of host interfaces.

The *HIP SM module* implements the main functions of the HIP State Machine. In my model transitions of HIP State Machine assume that packets are successfully authenticated and processed. This behavior is in consistence with the standards, therefore the skeleton implementation of security algorithms do not hamper the model to accurately simulate HIP mechanisms. One instance of *HIP SM* represents and manages one HIP connection with one Security Association. *HIP SM* handles transitions occurring during HIP Base Exchange, RVS registration, UPDATE mechanism, etc., and generates HIP messages according to the state transitions. *HIP SM module* also handles changes in partner IP addresses (sets the locators by receiving and processing UPDATE messages), but the actual storage happens in the main HIP module's hitToIpMap structure.

The *RvsHIP* and *LRVSHIP modules* are derived from the HIP module in order to extend the basic HIP capabilities with the RVS/LRVS functions by handling the incoming registration messages according to [78], by forwarding I1 messages [23] to the appropriate HIP responder chosen from the registered ones, and by implementing my proposed LRVS functions [C4], [C11], [C21], [B3], [J20].

The *DnsBase module* is a simple UDP application which realizes basic DNS server functionality for name resolution of HIP hosts and implements the new Resource Record (DNS HIP RR) defined in [84]. The module resolves domain names to HITs and IP addresses and in case of mobile HIP hosts also provides RVS information. Note, that reverse DNS lookups are not supported in the current version of HIPSim++.

With the help of the above modules, novel INET nodes can be created, which then can be placed into the appropriate simulation topology. HIP RFCs and Internet Drafts define three main types of nodes, namely the Initiator, the Responder and the Rendezvous Server. For introducing name resolution functions, also DNS server entity is to be used in a HIP architecture. All the above HIP nodes have been realized in HIPSim++ based on the existing INET modules [87] and the newly introduced *HIP*, *HIPSM*, *RvsHIP*, *LrvsHIP*, and *DNSBase* modules.

Novel message constructions were also required to be introduced in INET during the implementation of the HIPSim++ HIP simulation framework. HIP signaling messages, HIP user data messages, and DNS messages were created.

In accordance to [20], different *HIP signaling messages* start with a fixed header. The HIP header is logically an IPv6 extension header such in HIPSim++ all HIP messages are implemented as additions to the INET's *Ipv6ExtensionHeader*. Almost all the already

standardized HIP message types and parameters are defined in the framework, including also the Locator parameter which is realized as an array of HIPLocator structures. An important exception is the ESP_INFO parameter which is missing due to the simplified management of IPSec SPIs in my simulation model.

In HIPSim++ the Encapsulated Security Payload (ESP) based mechanism for transmission of *HIP user data messages* is applied [83]. As proper implementation of all the cryptographic mechanisms in HIP is outside of the scope of my researches, I use only simplified Encapsulating Security Payload Header [88] mechanisms for distinguish HIP data packets based on SPIs. Every HIP data message travels in ESP: packets coming from the transport layer will be encapsulated in an *ESPHeaderMessage* labeled with the appropriate SPI value. Every *ESPHeaderMessage* has a special object (called *IPv6EncapsulatingSecurityPayloadHeader*) per header to carry the SPI value as parameter. This object is derived from the *IPv6ExtensionHeader* class of INET in order to overcome some inflexibility issues of the existing IPv6 implementation and making the ESP packets to pass through the networking layer towards the HIP module.

The basic HIP namespace resolution functions are implemented using a simple query/response message pair called *DnsQuery* and *DnsResponse*.

In order to assess the standard parts of the above introduced model and to evaluate the accuracy and preciseness of the implementation, me and my co-authors in [C17] built and configured a real-life HIP testing environment based on InfraHIP [89], and compared the outcomes of the simulation with the reference results obtained from the testbed. Our analysis showed apparent accuracy and consistent operation of HIPSim++ in terms of handover metrics (latency, packet loss, throughput) and behavior in case of the standard HIP mobility mechanisms when compared to the experiences gathered in the real-life HIP testbed. This accuracy has been provided by modeling HIP messages, nodes and mechanisms in HIPSim++ based on the actual recommendations of current IETF RFCs, and by re-using the existing detailed IPv6, mobility, channel, etc. models of the INET Framework. This proved accuracy and degree of reliability of HIPSim++ makes the model ideal for evaluation efforts of ongoing and future HIP extensions, like my μHIP scheme.

2.2.3.2 Simulation scenarios and measurement results

I have used the standard HIP scenario as a reference, where the mobile HIP host (MN) changed its network point of attachment by connecting to another Wi-Fi access point (AP) due to its movement (Fig. 15 left). As the APs were connected to different access routers advertising different IPv6 prefixes, the IPv6 address of the MN was changed after reattachment. Standard HIP mechanisms were applied to handle this mobility situation by running the HIP UPDATE process [22]. During the simulation built-in TCP and UDP application models were used to generate traffic between the MN and its Correspondent Node (CN). I have introduced a special router node providing an average RTT of 300 ms between the MN and the CN / HIP RVS to simulate Internet-wide communication. A simple Domain Name Service model was applied used to simulate DNS procedures, but they were initiated only before connection establishment (i.e., HIP BEX).

For the μHIP scenario the difference lies in the introduction of micro-mobility domains: two HIP LRVSs replace the access routers and control their Domains (1 and 2), where the first one owns two access points (AP1, AP2) providing possibilities to simulate intra-domain handovers within its LRVS control node (Fig. 15 right). Inter-domain handovers are also implemented in the scenario: during its movement the MN changes its network point of attachment from AP2 to AP3 (belonging to Domain 1 and 2 respectively).

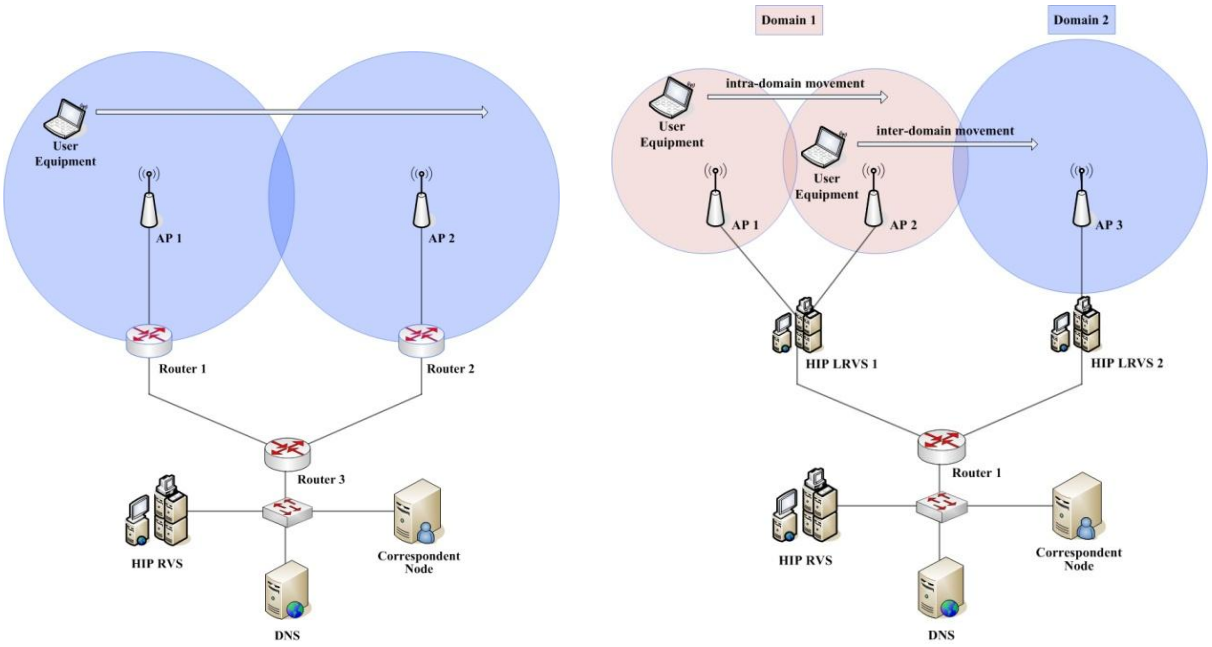


Figure 15: Simulation scenarios for standard HIP mobility (left) and my µHIP (right) scheme

In all the above scenarios the MN is able to migrate between the different APs with a constant speed such provoking handovers situations. By inducing 100 independent handovers during simulation runs I have measured three key performance indicators in three different sub-scenarios. *Sub-scenario A* measures Handover Latency defined here as the time elapsed between loosing the connection at the old AP and the MN sending out the last mobility management related signalling packet (e.g., HIP UPDATE packet) while connected to the new AP. The measurements were analyzed in function of different average Router Advertisement (RA) intervals such creating the simulation runs (each executed 100 times) defined by the different RA values. Fig. 16 presents the handover latency as the average of 100 independent handover series for every RA interval. I have shown that the latency of µHIP intra-domain handovers is approx. 10% better compared to the standard HIP performance. The much rarely occurring inter-domain cases produce approx. 6% higher values due to the additional management tasks when entering a new micro-mobility domain.

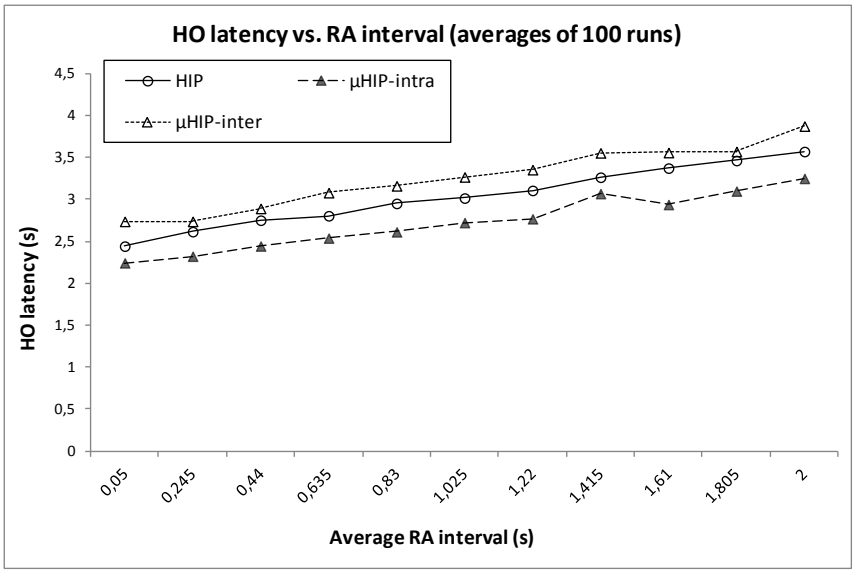


Figure 16: Handover latency measurement results of the µHIP scheme

Sub-scenario B measures UDP packet loss during handover situations in function of different data rates of the UDP application (by setting the inter-packet departure time) originated by the MN towards the CN. Fig. 17 shows how many UDP packets were lost during a handover in a HIP and µHIP based system. The points on the graph represent the average UDP packet loss of 100 independent handovers for every offered datarate value. The simulations clearly illustrate how µHIP enhances the handovers in intra-domain scenarios by the cost of slightly worse results for inter-domain HO events.

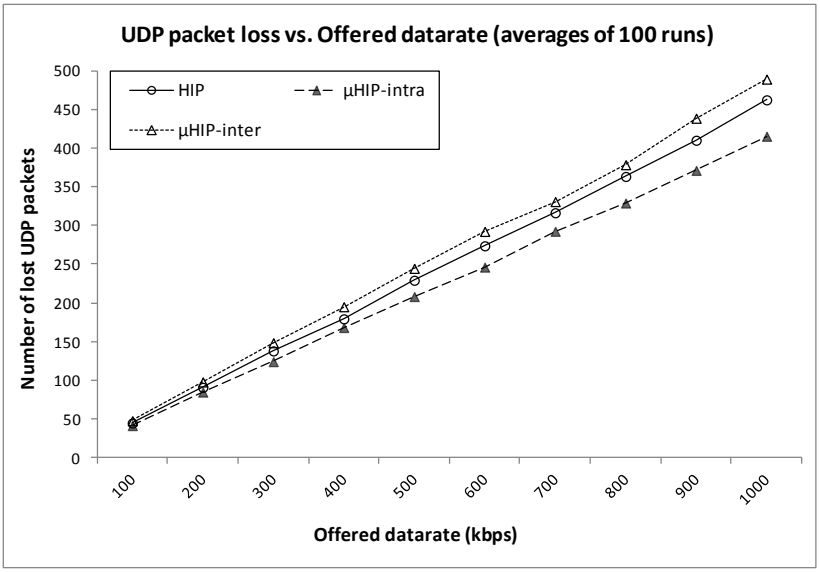


Figure 17: UDP packet loss measurement results of the µHIP scheme

In *sub-scenario C* I have measured TCP throughput of one minute experienced at different handover frequencies. Here the simulation runs were defining the number of handovers suffered by the MN per minute. Fig. 18 depicts the TCP throughput proportion in a one minute communication session between the MN and the CN experienced at different handover frequencies from 0 to 9. The gain of µHIP in intra-domain use-cases is 20% in average with the price of 9% decrease during the much less frequent domain changing situations.

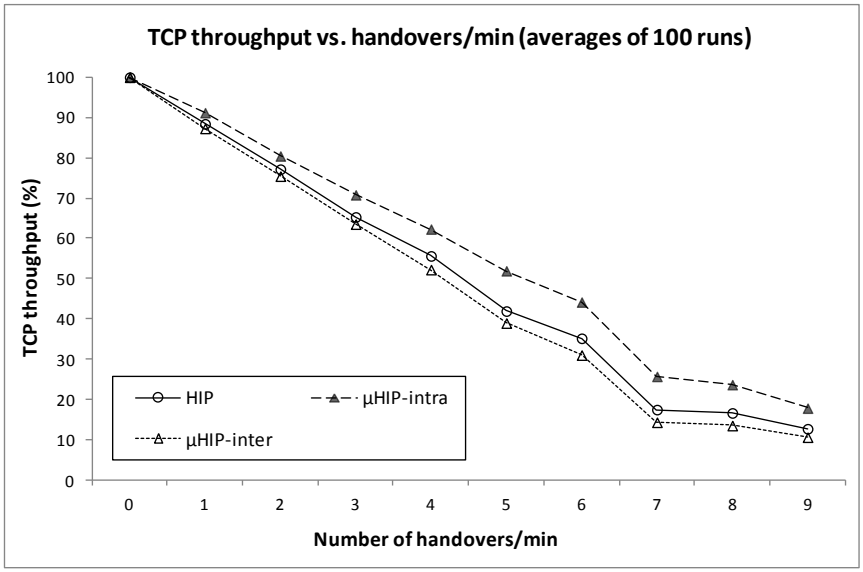


Figure 18: TCP throughput measurement results of the µHIP scheme

As a summary I can state that my HIP micromobility extension reduces signaling overhead and handover latency when it is used as a complement to the base HIP mobility management mechanism, thanks to its below summarized two main characteristics.

Update procedure: If a MN that uses normal HIP mobility function changes any point of attachment of the network, it must report its new IP address to the RVS and to all of the CNs throughout of the whole network. If my micromobility extension is used, and the MN moves to another subnet within the same domain, the only update procedure, which must be proceed, is to update the local LRVS. Since the LRVS always knows the correct in-domain IP address of the MN, the data packets sent by CNs to the MN are forwarded to this IP address by the LRVS. Thus there is no need to update the CNs or the RVS. This reduces the number of signaling messages correspondent to the location update, if MNs move frequently within a single domain. However, if the MN changes the visited domain, which is managed by another LRVS, the MN has to perform a registration mechanism with the new LRVS. This would not be necessary if standard HIP is used. The MN has to update all of its CNs (i.e., to report its new globally routable IP address to the CNs). This causes a slight overhead compared to the standard HIP solutions. Of course inter-domain movements are much more rare compared to intra-domain mobility events.

Handover process: If base HIP mobility function is used, and the MN changes its attachment point, it has to update all of the active HIP associations. Until a CN is not updated and sends data packets to the old IP address of the MN, these packets are lost. The more CNs have to be updated, the higher probability of packet loss is identified. In my method if the MN changes the visited domain, it does not immediately close its rendezvous association with the old LRVS; moreover, the MN updates its registration at the old LRVS by reporting its new globally routable IP address. This reduces the probability of potential packet losses during the handoff mechanism, since the MN is still reachable via the old LRVS.

Chapter 3
Location Privacy Aware Micromobility Domain Planning Schemes

Mobile terminals' location data possess important service-enabler potential, but in wrong hands it can be used to build up private and intimate profile of the mobile user and can pose serious threats to location privacy [90]. In the all-IP world of future mobile Internet, location privacy of users is even harder to protect as the most common parameters in every single packet – i.e., the source and destination IP addresses – can easily be translated to a quite accurate estimation of the peers' actual geographical location [91]–[95] thus making third parties able to track mobiles' real-life movements [96], [97]. In next generation all-IP heterogeneous wireless communication systems moving across multiple IP subnets (i.e., changing access nodes which provide IP connection with topologically different address spaces) will occur more likely, resulting in much frequent IP address changes compared to today's mainly homogeneous architectures, therefore further aggravate problems of location information leakage.

Micromobility solutions – besides localizing mobility events by grouping several IP subnets into domains, and providing fast, near-seamless and local handoff – also include capabilities to support location privacy: localization of mobility events inside a micromobility domain can hide location information easily exposable by IP address changes of handovers [79]. The reason is that only in cases of inter-domain handovers the location is updated and revealed to outside of the domain; not on each access point change. Pico- and femto-cell based mobile architectures [98] are even more sensitive to the above Quality of Service (QoS) and privacy issues which imply the spreading of micromobility protocols and the need of advanced network planning algorithms to support real-life deployment of the micromobilty paradigm.

As mobility becomes one of the most unique characteristics of future's convergent architectures, more attention must be paid to the location privacy issues, even at the earliest phases of design: at the network planning level. One of the open issues of deploying micromobility protocols in next generation mobile environments is the optimal design of domains. The main question is what size (in terms of consisting subnets) a micromobility domain should be for reducing the cost of paging, maintaining routing tables and registration signaling. Existing network planning algorithms (e.g., [24]–[27], [75], [J3]) are mainly focusing on the trade-off between the paging cost and the registration cost and – to the best of my knowledge – none of them have introduced privacy awareness in network planning methodologies. Also the potential of micromobility protocols to efficiently support location privacy was never taken into consideration in any domain planning algorithms available in the literature. This motivated me to develop mobile network planning tools that exploit inherent location privacy support of micromobility protocols while also considering the strict constraints formed by paging and registration costs (Thesis II.1, II.2, II.3 and II.4 in Sections 3.1.1, 3.1.2, and 3.2.2).

3.1. Privacy Aware Simulated Annealing based Location Area Forming

3.1.1 The proposed privacy model and algorithm

However there exists a quite broad literature on location area and micromobility domain planning as I introduced in Section 2.1.3.1, the substantial and a-priori question of how to integrate location privacy requirements into the algorithms is still almost completely

unexplored. To the best of my knowledge, the only study about location privacy aware domain planning was performed by me, firstly extending my SABAS solution with a simple location privacy policy model and a special rate weighting technique applied to integrate the effects of the cells' static location privacy significance and mobile nodes' dynamic privacy demands into the boundary crossing rates between neighboring cells. The algorithm is called PA-SABLAF (Privacy Aware SABLAF).

Thesis II.1. [J8], [B4] *I have developed a simple location privacy policy model to provide boundary conditions for location privacy aware domain planning where both static requirements and dynamic demands are to be respected. Based on this model I have proposed a special rate weighting technique for enhanced and privacy aware graph representation of mobile networks. Using this novel toolset I have developed a privacy aware domain planning algorithm called PA-SABLAF (Privacy Aware Simulated Annealing based Location Area Forming) which is an improvement of my SABAS algorithm decreasing the number of inter-domain handovers while also considering the location privacy in the created domain structure.*

In the location privacy policy model I have proposed, a combination of two substances is used to provide boundary conditions for location privacy aware domain planning. On the one hand I introduced the *static location privacy significance level of the cells* (denoted by $SLP_{[k]}$ for cell k) which can separate coverage areas inside the operator's network that are considered to be more sensitive to location privacy than others. On the other hand I defined *user's location privacy profile for different location types* (denoted by $ULP_u^{lt_{[k]}}$ for user u and location type lt of cell k) to describe what level of location privacy protection is required for a mobile user at a given type of location. The incoming dynamic demands are cumulated and the average will be compared with the static location privacy significance level of the issued cell at every announcement. The winner of this comparison – called the *cell's overall location privacy factor* – will take over the role of the cell's static significance level. In this simple way not only operators' requirements, but also the dynamic demands of mobile users can be respected during the location privacy aware network design.

In order to integrate the effects of the cells' overall location privacy factor into the boundary crossing rates between neighboring cells, I have created a special rate weighting technique. In the mathematical representation I applied, the cells are the nodes of a graph, the cell border crossing directions are represented by the graph edges and the weights are assigned to the edges based on the cell border crossing rates of every direction (i.e., rates of entering or leaving a cell are summarized and assigned to the corresponding edge as its weight). These rates are weighted with the overall location privacy factor of the destination cell.

$$WR_{[k][l]} = CR_{[k][l]} \times OLPF_{[l]} + CR_{[l][k]} \times OLPF_{[k]} \qquad (17)$$

where $WR_{[k][l]}$ is the weighted rate of edge between cells (graph nodes) k and l, notation $CR_{[k][l]}$ stands for the cell border crossing rate from cell k to l, and $OLPF_{[l]}$ is the overall location privacy factor of cell l.

Based on the above definition, my proposed PA-SABLAF algorithm starts with a GREAL-based greedy phase that will provide a basic domain partitioning as an input (i.e., initial solution) of the simulated annealing. At the beginning of this greedy phase, we choose the cell pair with the biggest weighted rate in our cell structure. If the biggest rate occurs multiple times, then we choose one of the instances randomly and include the two cells

belonging to that handover rate into domain D_1 of cells. In the next step, we search for the second biggest weighted rate among the cell pairs for which is true, that one of them belongs to domain D_1. We must check whether inequality $N_k < N_{max}$ is satisfied, where N_k is the number of cells in the k^{th} domain and N_{max} stands for the maximum number of cells in a single micromobility domain which will give us the minimum of the registration cost and the maximum size of the location privacy protective micromobility domain. If the inequality is satisfied, the cell can be included into set D_1. If the inequality is not satisfied, the cell can not be included into this set: a new domain with this cell is to be created in order to prevent exceeding the paging cost constraint. In this way we can join the most important cells according to the location privacy policy model which are also in the same dominant moving directions (highways, footpaths, etc.,).

After processing all the cell pairs in the above sequential and greedy way a likely sub-optimal domain structure will be created, which will serve as an input (i.e., initial solution or s_0 domain partitioning) for the simulated annealing part of the algorithm. Based on s_0 a neighbor solution s_1 is then generated as the next solution ($N_k < N_{max}$ must be satisfied by s_1 too), and the change in the registration cost $\Delta C_{Reg}(s_0, s_1)$ is calculated. If a reduction in the cost is achieved, the current solution is replaced by the generated neighbor; otherwise we evaluate the acceptance function $e^{(-\frac{\Delta C_{Reg}}{T})}$ to decide whether to retain or change the current solution (T is the temperature). The cooling schedule is based on three input parameters: initial temperature T, step of decrement ($decr$) for T, and the stopping rule which is the maximal iteration step number until ΔC_{Reg} does not change.

3.1.2 Initial metric and evaluation

In order to evaluate the developed domain planning scheme, I have designed a privacy metric and extended the simulation framework introduced in Section 2.1.3.7.

Thesis II.2 [J8], [B4] *I have proposed a location privacy metric called* $\boldsymbol{LP^s_{mic}}$ *to express how efficiently a given micromobility domain structure takes into account static location privacy significance of cells and the incoming dynamic location privacy demands of users during operation. I have shown that PA-SABLAF appreciably improves the domain structure compared to its predecessor algorithm with an average of 10% location privacy gain.*

I have proposed $\boldsymbol{LP^s_{mic}}$ to show how effective could be the protection of users' location privacy while keeping paging and registration costs on a bearable level in a given micromobility environment. I have quantified the inability of non inside-domain attackers in tracking mobile users by computing a weighted number of inter-domain changes of mobile nodes in the network. This metric tracks and saves movements (i.e., whole paths) of mobile users and also saves cell boundary crossings in order to localize and count mobile nodes' inter-domain changes. For every inter-domain handover of a mobile node and for the previous and the next cells of such handovers the metric calculation algorithm sums the value of the cells' static location privacy significance and the squared value of the level of the mobile node's location privacy profile set for the issued location types. The above calculation is performed for every mobile node, and the sum of these values will stand for the location privacy metric of the whole micromobility domain system:

$$\boldsymbol{LP^{s}_{mic}} = \sum_{u} \sum_{h \in IH_u} (ULP_u^{lt_{[k]}})^2 + (ULP_u^{lt_{[l]}})^2 + SLP_{[k]} + SLP_{[l]} \tag{18}$$

where IH_u means the set of all inter-domain handover events of user u, and $h_{[k][l]} \in IH_u$ stands for a handover event with exit and entry cells of k and l respectively. Implicitly the smaller $\boldsymbol{LP^{s}_{mic}}$ values are the better.

3.1.2.1 Simulation environment and evaluation results

I have evaluated PA-SABLAF in a further extended version of the mobile environment simulator already introduced in Thesis I.2 (Section 2.1.3.7). The performed enhancements on this Java-based system are the followings.

1. The system also calculates both the handover rate and the location privacy-weighted rate for each cell/PoA pair, defined on the border of these cells/PoA coverages.
2. The static location privacy significance level of the cells can also be set in case of need as well as the location type.
3. Mobile nodes can be placed into this highly customizable environment by firstly specifying MNs' velocities, setting the incoming session arrival parameter (IP session intensity) and also the location privacy profile to every mobile node if needed.
4. Different types of mobility environments with different location privacy characteristics can be modeled (rural environment with highways without strict location privacy requirements or a densely populated urban environment with roads and carriageways and the widest scale of location privacy sensitive areas like military facilities, government buildings, etc.,), together with the grids of cells configured and adapted to these environments.
5. When a simulation run ends, the simulator sums the cell boundary crossings and incoming session initiation distribution for every cell in the simulated network, and also calculates the normal and the location privacy-weighted rates for the micromobility domain planning algorithms.

My goal with these extensions was to provide a more flexible tool which is able to give the possibility to evaluate LA partitioning and micromobility domain planning algorithms for the widest scale of network types, by freely choosing the road grid, communicating mobile hosts and cell structure/characteristics.

The evaluation was carried out with the help of two key performance indicators. On the one hand I analyzed PA-SABLAF using the applicable privacy metric ($\boldsymbol{LP^{s}_{mic}}$) from the location privacy point of view. On the other hand I used the global registration cost to measure the efficiency of the algorithm from the signaling cost optimization perspective. Note, that besides the above I also considered the N_{max} as a constraint for the paging costs.

I have executed several simulation runs for PA-SABLAF for $N_{max} = [2\,..\,6]$ values, for every scenario, and depicted the total average of all the measurements for a particular domain planning solution in function of the N_{max}.

Four different scenarios were defined and created in this simulation framework by cell/PoA, mobile node and movement path placing. These scenarios were designed to differ in their cell/access point structures, number of active mobile users, and style of interconnection (i.e., possible transition paths between cells) aiming to provide a reasonable scale and variety of initial input data for evaluation. Fig. 19 depicts these scenarios and the following enumeration details the most important scenario parameters and characteristics.

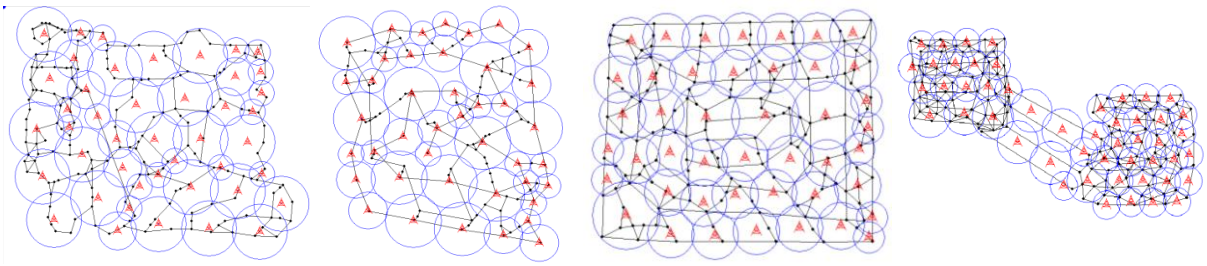


Figure 19: Simulation scenarios used for evaluation (#1, #2, #3, #4 from left to right, respectively)

1. *Scenario #1* consists of 44 multiply interconnected cells and 33 mobile users. Both densely linked (urban-like) and rarely linked (rural-like) areas exist in this construction.
2. *Scenario #2* consists of 42 multiply interconnected cells and 33 mobile users. The average level of interconnection of cells is significantly lower than in Scenario #2, implicating smaller number of transition possibilities.
3. *Scenario #3* consists of 44 multiply interconnected cells and 25 mobile users. This scenario represents a structure where the possible number of inter-cell transitions is high.
4. *Scenario #4* consists of 50 multiply interconnected cells and 22 mobile users. This structure has two, densely linked cell groups, which is interconnected with only a limited set of cells and transition paths.

The simulation of the four scenarios was run till the completion of several thousands of handovers in order to generate substantial number of realistic cell boundary crossings, incoming call/session data and location privacy-weighted rates for each cell pair, also calculating the paging cost and the registration cost for every domain. The produced data will then be used as an input for the algorithms to be evaluated.

Using this environment I have compared my algorithm with its ancestor – the already introduced SABAS which is without any trace of location privacy awareness. As an initialization of my experiments I ran the mobility simulator on the scenarios of Fig. 19 and gathered all the required input data for PA-SABLAF and for the base algorithm of the evaluation (i.e., SABAS). After that I executed all the algorithms (with parameters $N_{max} = 6$, $T = 100$, and $decr = 2$) on the produced input data and cell structure in order to render the micromobility domain configuration. On the rendered domain layout I examined how the registration cost and the location privacy metric changes by increasing the maximum number of cells in one micromobility domain for each algorithm and scenario. This way I could check how the domain forming methods perform in terms of location privacy support and signaling cost optimization, and also whether the registration cost function is correct (i.e., whether it reaches the minimum value when a domain consists N_{max} number of cells.)

Simulation results show (Fig. 20) that PA-SABLAF finds a much better domain structure in terms of the $\boldsymbol{LP^{s}_{mic}}$ metric for every value of N_{max} compared to the original SABAS. However, we have to pay the price of this benefit: the registration cost is slightly higher in most of the cases with a maximum of 4.8%. I have to emphasize, that the $N_{max} = 6$ case results in gain also regarding the registration cost, so here the algorithm managed to ameliorate both parameters of the trade-off.

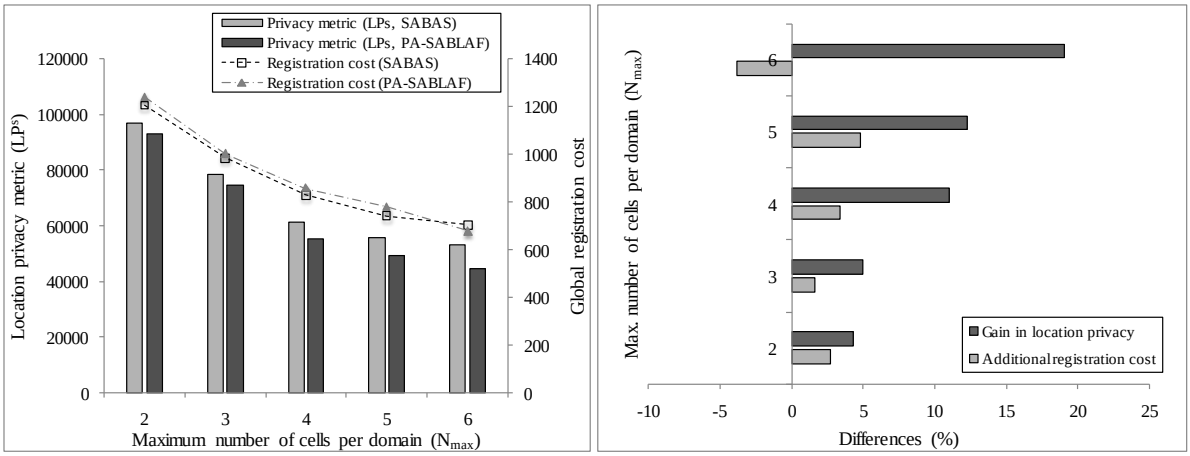


Figure 20: PA-SABLAF vs. SABAS (left) and Location privacy gain vs. cost increment for PA-SABLAF (right)

3.2. Adaptation and application of existing location privacy metrics to domain planning

However $\boldsymbol{LP^{s}_{mic}}$ is able to numerically present the location privacy capabilities of a complete network's certain micromobility domain structure, it lacks in generality. That is why I have started to evaluate my scheme using more general and widespread location privacy metrics.

3.2.1 Introduction to existing location privacy metrics

3.2.1.1 Uncertainty-based location privacy metric

This type of metric was originally proposed in [99], [100]. Authors in [99] present an information theoretic model that allows to measure the degree of anonymity provided by schemes for anonymous connections. Authors of [100] introduce an information theoretic measure of anonymity that considers the probabilities of users sending and receiving the messages and also show how to calculate this measure for a message in a standard mix-based anonymity system. Both proposals use the same metric model. The attacker's goal is to identify the initiator and/or the responder of a message travelling in the network. Each user in the system is assigned a probability for being the possible initiator/responder of a particular message, and the system's overall anonymity level is determined by the entropy of the random variable that is formed by the users' probabilities. In this way the metric captures the attacker's uncertainty (measured by the entropy) during the identification procedure.

This metric can be easily applied to comply also with location privacy measurement purposes: the location privacy of a given user in the system is calculated as the attacker's uncertainty during linking observed events (e.g., positions in trajectories) to users. Authors in [101] define a well-detailed system model which can be used to formalize the uncertainty-based location privacy metric as follows.

Consider a user $u \in U$ and an event $\hat{e}_i \in \hat{R}_u|_{obs}$ successfully observed by the attacker ($\hat{R}_u|_{obs}$ means all the user u events observed by the attacker). Also consider E^l as the set of edges of the probabilistic graph called the linkability graph (G^l) representing the linkability of observed events based on the attacker's knowledge. (Note, that the attacker's goal is to reassemble the user's actual set of events; hence it assigns probabilities to possible related

events in order to reconstruct the user's trajectory.) Define $\pi^l(\hat{e}_i, \hat{e}_j)$ as the weight function of G^l for representing the probability with which the attacker believes that both $\hat{e}_i$ and $\hat{e}_j$ events are associated with the same user and $\hat{e}_i$ is an immediate predecessor of $\hat{e}_j$ in the user's observed set of events. Let a random variable X_i stand for the probability that another observed event $\hat{e}_j$ is the immediate successor of $\hat{e}_i$. According to the notations of [101] we have $Pr_A(X_i = j) = \pi^l(\hat{e}_i, \hat{e}_j)$ for any j such that $(\hat{e}_i, \hat{e}_j) \in E^l$ in G^l, where $Pr_A(s)$ means the probability with which the attacker considers a statement s to be true. The entropy of X_i can be calculated to measure the attacker's uncertainty when linking observed events to users and therefore can be used as objective location privacy metric for user u at the $tm(\hat{e}_i)$ time instance at which the event $\hat{e}_i$ occurred.

$$\boldsymbol{LP_u^u}\big(tm(\hat{e_i})\big) = \mathbb{H}(X_i)$$
$$\mathbb{H}(X_i) = -\sum_j Pr_A(X_i = j) \times log_2\big(Pr_A(X_i = j)\big) \tag{19}$$

Eq. (19) is the common form of the uncertainty-based metric which has been widely used to measure location privacy in different wireless scenarios. For example, authors of [102] applied this scheme in order to maximize the location privacy at each identifier update by users, in the presence of asynchronous identifier updates and predictability of movements of user terminals. In [103] this metric is used to analyze a novel location privacy enhancement protocol which obfuscates several types of privacy-compromising information revealed by mobile nodes, including sender identity, time of transmission, and signal strength. The last example for application use-cases of the uncertainty-based metric is [104] where authors employ the scheme to measure the performance of their solution designed to enhance users' contradictory requirements on location privacy without diminishing communication QoS.

These examples show how varied application possibilities of the uncertainty-based location privacy metric are and why it is considered a widely accepted and adapted metric in the literature.

3.2.1.2 Traceability-based location privacy metric

This kind of metric captures the level to which the attacker can track a mobile user with high certainty. The attacker's uncertainty or confusion in the tracking procedure is measured by the uncertainty-based metric (e.g., using entropy). In [105] authors define a so called *mean time to confusion* metric to measure the degree of privacy as the time that an attacker could correctly follow a user's trace. Therefore the *mean time to confusion* is the mean tracking time between points where the attacker faced confusion (i.e., was not able to determine the next sample with sufficient certainty). Authors of [106] propose another variant of the traceability-based location privacy metric called the *mean distance to confusion*, which measures the mean distance over which tracking of a user may be possible by the attacker.

The above variants are defined as the time/travel distance until the uncertainty of the tracking grows above a pre-defined threshold. The formalization of the traceability-based location privacy metric is also done according to the system model defined in [101]. They call an event in the observed trace of a user ($\hat{e} \in \hat{R}_u|_{obs}$) as a confusion point if the attacker's uncertainty is above a given threshold: $\boldsymbol{LP_u^u}(tm(\hat{e})) > \mathbb{H}_{cf}$. It means that the *time to confusion / distance to confusion* is specified as the time/distance to travel before reaching a confusion point, during which the attacker's uncertainty remains below $\mathbb{H}_{cf}$. In this way the average value of *time/distance to confusion* represents a user's lack of location privacy.

Denote $\hat{R}_u|_{obs}^{cf}$ as the set of all confusion points (events) of user $u \in U$. Let C_u stand for

the union set of the last observed event of user u and the user's confusion events. Let B_u denote the set of events that contain the first observed event from u and all the events which are not confusion points but are immediate successors of each confusion point in the observed trace of user u. Consequently, a traceable period can be defined as the time/travelled distance between an event in B_u and an event in C_u such that there is no other event in B_u in that period. Denote Z_u as the set of all these traceable periods for user u. Based on the above notation the location privacy metric of user u based on *mean time to confusion* ($\boldsymbol{LP}_u^{\dot{t}}$) and *mean distance to confusion* ($\boldsymbol{LP}_u^{\ddot{t}}$) can be defined as the mean tracking time/distance during which uncertainty stays below a confusion threshold and can be calculated as follows (note, that this metric is inversely proportional to *mean time to confusion* and *mean distance to confusion* values).

$$\boldsymbol{LP}_u^{\dot{t}} = \left(\frac{\sum_{(\hat{e}_i, \hat{e}_j \in Z_u)} \left| tm(\hat{e}_i) - tm(\hat{e}_j) \right|}{|Z_u|} \right)^{-1} \tag{20}$$

$$\boldsymbol{LP}_u^{\ddot{t}} = \left(\frac{\sum_{(\hat{e}_i, \hat{e}_j \in Z_u)} \left\| loc(\hat{e}_i) - loc(\hat{e}_j) \right\|}{|Z_u|} \right)^{-1} \tag{21}$$

where $tm(\hat{e}_i)$ and $loc(\hat{e}_i)$ stands for the time instance and the location at which the event $\hat{e}_i$ occurred, respectively.

This metric is also a well-known and widespread measure of location privacy. One of its main advantages is the fine-grained tuneability: if the threshold of $\mathbb{H}_{cf}$ is chosen high, tracking times increase but so may do the number of false positives (i.e., the attacker follows incorrect traces). A good example for the application of this metric is [107] where authors used the approach in a trace-based simulation of their anonymizer scheme camouflaging users' current location with various predicted paths.

Both the above introduced uncertainty- and traceability-based location privacy metrics are general, widespread and also effective in the means that they are able to quantify the incapacity of a particular attacker in localizing or tracking mobile users. That was my main motivation for choosing them as base approaches in my evaluation work and efforts to further enhance PA-SABLAF.

3.2.2 Realization/adaptation of the metrics and improving PA-SABLAF

3.2.2.1 Metric requirements and assumptions

The level of location privacy in a complete network (i.e., system of several micromobility domains) can be determined by how easily attackers can recognize trajectories (series of cells/access areas owning unique IP prefixes) of mobile users. Every single user in the mobile network passes cells of several domains during their respective paths.

In such architecture inside-domain movements are safe as localized mobility management obfuscates IP address changes of mobile users: valuable addressing information (i.e., location information data of IP communication) will not leak out the domain. However, domain changes will disclose IP address information to all correspondent nodes (CNs) of the mobile as macromobility mechanisms will also be executed besides the micromobility procedures. I assume that the attacker is in continuous communication with the observed mobile user (or at least the attacker is able to capture packets originating from the MN) and is located outside of

the MN's domain. The obtainable address information is enough to identify the MN's actual domain but not sufficient to determine the particular cell or access area from where the mobile node communicates.

Due to the aforementioned characteristics of the micromobility management the attacker continuously communicating with the MN can be aware of the complete set of domains crossed during the MN's path, and this information can be used to specify the precise, cell-based trajectory of the observed entity (i.e., the solution which the attacker wants to obtain). Reconstruction of the whole trajectory gets harder if the built-in location privacy supporting capability of the micromobility domain system (i.e., the obfuscation of the observable information performed by the localized mobility management) becomes more effective. This is what a metric in my framework proposal should measure.

As the attacker can observe only the set of domains the MN passes, it must apply statistical calculations to get the solution. An adequately large domain with sufficient number of inter-domain transitions is able to significantly increase the quantity of potential solutions and to enhance location privacy of users in a general way, independently of pre-defined or dynamic privacy parameters I applied in Section 3. The metrics below serve as efficient tools in our efforts to enhance PA-SABLAF and create more comprehensive and universal algorithms.

3.2.2.2 Realization/adaptation of $\boldsymbol{LP^u}$ and introduction of the PA^u-SABLAF variant

Aiming to implement the uncertainty-based metric in my simulation framework and to adapt it for evaluation purposes I have slightly modified the original $\boldsymbol{LP^u}$ scheme. In order to do this, I adapt the $\boldsymbol{LP^u}$ scheme to my framework and also extend it to be applicable for all the users in the micromobility system (as a sum of their entropies).

Thesis II.3 [J8], [B4] *I have proposed a location privacy metric called $\boldsymbol{LP^u_{mic}}$ in order to adapt the uncertainty-based location privacy metric for localized mobility scenarios and measure the level of obfuscation provided by the built-in location privacy supporting capability of micromobility domain systems. I have developed a privacy aware domain planning algorithm variant called PA^u-SABLAF to enhance the domain planning process in terms of the $\boldsymbol{LP^u_{mic}}$ metric. I have shown that PA^u-SABLAF is able to improve the domain structure with a significant 30% relative growth in high PoA number domains by raising the possible number of transitions at inter-domain movements.*

In a micromobility network the attacker relying on intercepted IP packets can only observe series of crossed domains along the MN's movements. This is because domains usually contain several cells/PoAs with multiple possible transitions inside and outside of the particular domain. The exact place of a domain change (i.e., the two cells of that transition) can be determined with probability $\frac{1}{n}$ where n is the number of possible transitions between the two domains.

That is why I calculate $\boldsymbol{LP^u}$ in the following way. I split the trajectory of the MN into domain entry and exit points which are basically the observable events (locations) in our threat model and delimit unobservable path segments between them. As these inside-domain path segments are not traceable based on IP information and assuming that domains contain more than two cells at least, the attacker can only deduce the entry and exit points (so called "flashes"). I assume that transitions are not weighted and the transition probability is the same in every case. Considering $Pr_A(d)$ here as the probability of the attacker guessed right when reckoning the actual entry and exit points of crossing domain d, a user's $\boldsymbol{LP^u}$ for a particular

domain inside the network can be produced by calculating the entropy of $Pr_A(d)$. (Note that $Pr_A(d)$ can be computed as the product of the probabilities of inlet- and outlet routes belonging to two consecutive "flashes".) By calculating this entropy for every domain of every user, and creating the sum of these entropies we get the overall entropy of a micromobility system denoted by $\boldsymbol{LP^u_{mic}}$ (as this metric is an entropy-like measure, the larger values denote the better location privacy support).

$$\boldsymbol{LP^u_{mic}} = -\sum_u \sum_j Pr_A(d_j) \times log_2\left(Pr_A(d_j)\right) \tag{22}$$

In order to create a more general domain planning scheme based on the criteria of the widespread and universal uncertainty-based location privacy metric I have designed the PA^u-SABLAF algorithm variant. The main design choice for this algorithm was to eliminate the dependency of the operation from both the *static location privacy significance level of the cells* and the *mobile node's location privacy profile* (which equally can narrow the applicability of the model) and create a more general scheme based on the criteria of the widespread and universal uncertainty-based location privacy metric.

In order to do this I altered the greedy phase of the algorithm for increasing the uncertainty of the attacker during its tracking intentions by dismissing the privacy weighted boundary crossing rates ($WR_{[k][l]}$) and creating a novel weighting technique which raises the possible number of transitions at inter-domain movements.

For that reason the greedy phase of PA^u-SABLAF also considers the crossing rates of all the neighboring transitions besides the crossing rates of the actually examined transition. It means that during the contraction the greedy phase favors to choose cell pairs rendering big crossing rates and also showing big traffic through large number of edges between their neighbors. Since the maximum number of cells in a single micromobility domain is limited by N_{max}, we can always create a structure where cells with big transition rates will create domains and simultaneously their neighbors with reasonably significant number and volume of transitions will form neighboring domains thus increasing the uncertainty of the attacker observing users' domain changes. According to this, PA^u-SABLAF will lead the traffic of cells with large transit demands away toward as many edges/edge series as possible. The calculation of the weighted rate based on the above considerations and used in the greedy phase of PA^u-SABLAF is as follows.

$$WR^u_{[k][l]} = CR_{[k][l]} + CR_{[l][k]} + TF_{[l]} \tag{23}$$

where $CR_{[k][l]}$ stands for the cell border crossing rate from cell k to l, and $TF_{[l]}$ is the transition factor of cell l (a cell still waiting to be grouped into a domain). I defined the transition factor as $TF_{[l]} = \sum_{m \in A_l}(CR_{[l][m]} + CR_{[m][l]})$ where A_l means the set of all neighbors of cell l. Besides this modified weighting and cell selection scheme the PA^u-SABLAF algorithm is the same as the method introduced in Thesis II.1.

Using the simulation environment and parameters introduced in Section 3.1.2.1 and applying $\boldsymbol{LP^u_{mic}}$ in the framework I have shown that PA^u-SABLAF achieves serious relative gain in terms of the location privacy metric and the registration cost increment: a more then 30% relative growth can be noticed for location privacy in the $N_{max} = 6$ case (Fig. 21). Despite this promising result PA^u-SABLAF shows the most serious volume of additional registration costs after location privacy aware domain planning: even the smallest cost growth is 27%. However, this is compensated by the remarkable revenues of the $\boldsymbol{LP^u_{mic}}$ metric.

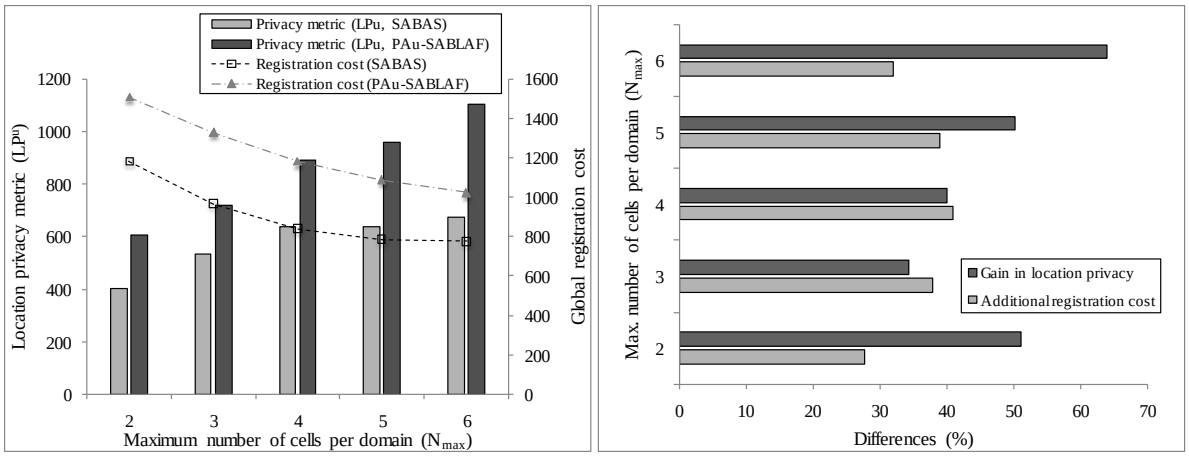


Figure 21: PA[u]-SABLAF vs. SABAS (left) and Location privacy gain vs. cost incr. for PA[u]-SABLAF (right)

3.2.2.3 Realization/adaptation of $\boldsymbol{LP^t}$ and introduction of the PA[t]-SABLAF variant

Due to the peculiar application scenario devised by my domain planning scheme, modifications in the original concept of the traceability-based metric ($\boldsymbol{LP^t}$) were required.

Thesis II.4 [J8], [B4] *I have proposed a location privacy metric called* $\boldsymbol{LP^{\bar{\bar{t}}}_{mic}}$ *in order to adapt the traceability-based location privacy metric for localized mobility scenarios and quantify the incapacity of attackers in localizing or tracking mobile nodes in a micromobility domain system. I have developed a privacy aware domain planning algorithm variant called PA[t]-SABLAF to enhance the domain planning process in terms of the* $\boldsymbol{LP^{\bar{\bar{t}}}_{mic}}$ *metric. I have shown that PA[t]-SABLAF is capable to improve the domain structure with an average gain of 3.9% by transacting and keeping user traffic inside the domains and also decreasing the registration cost in most of the cases.*

During the realization and adaptation phase of this kind of location privacy measurement approach I recognized that according to my scheme and threat model the attacker is not able to track mobile users when they are moving inside a particular micromobility domain. It means that domains serve as confusion points, which also implies that *mean time to confusion* and *mean distance to confusion* approaches become vague: users spend their time mostly in confusion points and only inter-domain handovers ("flashes") are considered as inter-confusion point events which are negligible both in terms of time and distance.

This motivated me to create two slightly modified traceability-based metrics called *mean time in confusion* and *mean distance in confusion*. These two metrics capture the level to which the attacker cannot track a mobile user with high certainty. The mobile user's safety during the IP information-based tracking procedure is measured by my modified $\boldsymbol{LP^t}$ metric versions. I define the *mean time in confusion* metric to measure the degree of privacy as the time that an attacker could not correctly follow a user's trace: the *mean time in confusion* is the mean tracking time between points where the attacker overcomes the confusion (i.e., becomes to be able to determine the next sample with sufficient certainty). Similarly, the *mean distance in confusion* measures the mean distance over which tracking of a user may not be possible by the attacker.

According to the already introduced formalization C_u stands for the union set of the last observed event of user u and the user's confusion events, B_u denotes the set of events that

contain the first observed event from u and all the events which are not confusion points but are immediate successors of each confusion point in the observed trace of user u. Consequently, an untraceable period can be defined as the time/travelled distance between two or more consecutive events in C_u such that there is no other event in B_u in that period. Let Y_u stand for the set of all these untraceable periods for user u. Based on the above notation the location privacy metric of user u based on *mean time in confusion* ($\boldsymbol{LP}_u^{\bar{t}}$) and *mean distance in confusion* ($\boldsymbol{LP}_u^{\bar{\bar{t}}}$) can be defined as follows.

$$\boldsymbol{LP}_u^{\bar{t}} = \left(\frac{\sum_{(\hat{e}_i, \hat{e}_j \in Y_u)} \left| tm(\hat{e}_i) - tm(\hat{e}_j) \right|}{|Y_u|} \right)^{-1} \tag{24}$$

$$\boldsymbol{LP}_u^{\bar{\bar{t}}} = \left(\frac{\sum_{(\hat{e}_i, \hat{e}_j \in Y_u)} \left\| loc(\hat{e}_i) - loc(\hat{e}_j) \right\|}{|Y_u|} \right)^{-1} \tag{25}$$

where $tm(\hat{e}_i)$ and $loc(\hat{e}_i)$ stands for the time instance and the location at which the event $\hat{e}_i$ occurred, respectively.

My simulation framework is not prepared for measuring the time duration between user events (i.e., handovers); the system fits only for marking locations (i.e., cells/PoAs) and the distance between different locations in terms of required transition numbers. Therefore I calculate the overall traceability-based location privacy metric of a micromobility system ($\boldsymbol{LP}_{\boldsymbol{mic}}^{\bar{\bar{t}}}$) in my simulator as follows (the location privacy supporting capability is proportional with the *mean distance in confusion*, so here the exponent implies that the smaller values are the better).

$$\boldsymbol{LP}_{\boldsymbol{mic}}^{\bar{\bar{t}}} = \sum\nolimits_u \left(\frac{\sum_{(\hat{e}_i, \hat{e}_j \in Y_u)} \left\| loc(\hat{e}_i) - loc(\hat{e}_j) \right\|}{|Y_u|} \right)^{-1} \tag{26}$$

The algorithm variant created based on the above $\boldsymbol{LP}_{\boldsymbol{mic}}^{\bar{\bar{t}}}$ metric also breaks with the rate weighting technique of my original PA-SABLAF and focuses on more general requirements characterized by the traceability-based location privacy metrics. Here the motivation is to create a micromobility domain structure where user traffic is mainly transacted and kept inside the domains. In case of PA$^{\text{t}}$-SABLAF I also approach this problem by modifying the applied weighting scheme of the greedy phase inside the original algorithm.

The traceability-based metric implies a single domain covering all the access areas (i.e., cells/PoAs) as the optimal solution for the location privacy aware domain planning problem. Of course this is not an option: N_{max} is the maximum number of cells in a single micromobility domain in order to provide a strict burden for the paging (and such also maximizing the size of the location privacy protective micromobility domain). So I have to take the cost constraints into consideration and simultaneously create a domain structure in which mobile users will likely perform inside-domain movements.

This can be achieved by increasing the number of "deflector" edges inside the domains. I define an edge or a series of edges as "deflector" if it possesses significant crossing rate and/or it provides input and output for high crossing rates of other edges or series of edges from multiple directions. By inserting cell pairs with deflector edges into the micromobility domains we can enforce that frequent cell/PoA sequences of mobile users will likely consist a domain. Such a structure decreases inter-domain movements while fulfilling all the domain

planning constraints and also enhances the privacy level of the micromobility scheme in an efficient manner. The calculation of the weighted rate based on the above introduced idea framed for the greedy phase of PA^t-SABLAF is as follows.

$$
\begin{aligned}
&if \\
&\qquad E_{[k][l]} \in D_{\psi} \\
&then\ for \\
&\qquad \forall E_{[i][j]} \in E_{[k][l]} \cup A_{[k][l]} \\
&do \\
&\qquad WR^{t}_{[i][j]} = CR_{[i][j]} + CR_{[i][j]} + DF
\end{aligned}
\tag{27}
$$

where $E_{[k][l]}$ denotes the edge between cells k and l, D_{ψ} means the set of deflector edges containing edges with the upper ψ percent of all crossing rates in the network, $A_{[k][l]}$ is the set of neighbors of $E_{[k][l]}$, $CR_{[k][l]}$ stands for the cell border crossing rate from cell k to l, and DF is a constant called deflector factor used for rewarding certain edges with deflector properties. This basically means that deflector edges chosen with parameter ψ and their neighboring edges are rewarded with parameter DF.

Besides the special weighting technique of (27) the PA^t-SABLAF algorithm is basically identical to my original scheme introduced in Thesis II.1.

The simulation results of PA^t-SABLAF evaluation are depicted in Fig. 22. The simulation environment and scenarios were the same as introduced in Section 3.1.2.1, but in this case I used the $\boldsymbol{LP^{\bar{\bar{t}}}_{mic}}$ metric and also applied different ψ and DF value combinations ($\psi = 20\%, DF = 20,\ \psi = 40\%, DF = 15, \psi = 60\%, DF = 10$), and showed the average of these results in my analysis. The PA^t-SABLAF algorithm variant performs a moderate average gain (3.9%) and also shows negative relative gain in the $N_{max} = 4$ case. However, the algorithm enhances the privacy metric together with registration cost in all the other N_{max} cases which is a valuable achievement.

As a result of my efforts in location aware micromobility domain planning I can state that the proposed scheme proved its power by significantly enhancing the location privacy of users in the network. The total average gain in location privacy for every run of all the three algorithm variants I developed approached 20% at the expense only of a total average 8% growth of the global registration cost (meaning an average 12% relative gain), and there were also distinct cases when the scheme operated with more than 30% relative gain.

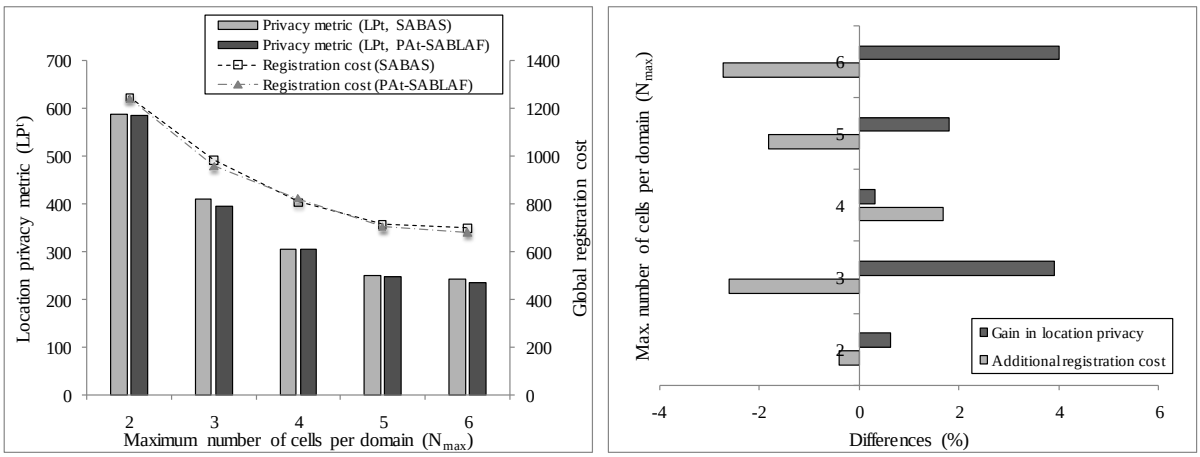


Figure 22: PA^t-SABLAF vs. SABAS (left) and Location privacy gain vs. cost incr. for PA^t-SABLAF (right)

Chapter 4
Optimized Solutions for Network Mobility Management

Trends in information technology show that heterogeneous, IP-based wireless networks will support mobility for the widest range of single end terminals (e.g., mobile phones, SmartPhones, PDAs, tablets and other handhelds), and even Personal Area Networks (PANs), Vehicle Area Networks (VANs) [12], Intelligent Transportation Systems (ITSs) and Cooperative ITS (C-ITS) architectures [13], [C5], [C10], [C15], networks of RFID (Radio Frequency Identification) devices and sensors, and various mobile ad hoc networks [14] will have permanent Internet connectivity during movement. Hence, in next generation wireless telecommunication not only single mobile entities have to be taken into account (host or terminal mobility), but also entire mobile networks moving between different subnets need to be maintained as a whole (i.e., network mobility or NEMO). IPv6 has introduced support for both mobility cases by defining Mobile IPv6 (MIPv6) [6] and Network Mobility Basic Support (NEMO BS) [15]. With these mobility supporting mechanisms all sessions remain active, even when the mobile node/network changes its subnetwork. When a host or a moving network has multiple interfaces and/or several IPv6 addresses, it is regarded multihomed. Multihomed mobile hosts/networks need special protocols to support their mobility management (e.g., MCoA [108], Flow Bindings [109], [110]). Handover at network layer usually takes several seconds due to the large number of L1/L2/L3 processes, the lack of interaction between them, and their complexity. The overall time needed to complete these procedures could go up to several seconds. In order to ensure seamless, continuous communication, this huge outage should be avoided by applying optimized handover solutions in the architecture.

Several improvement proposals exist to overcome the huge delay. All of them aim to speed-up the handover process. Mobile IPv6 Fast Handovers [111] is one example, and there are plenty of other proposals as well [36]–[38], [112], [113], [B7]. However, according to my best knowledge, none of the existing solutions exploit the benefits of overlapping radio access coverages by managing multiple tunnels and predictive tunnel switching. In order to enhance NEMO solutions, I have followed two approaches. On the one hand I have extended standard IPv6-based network mobility by forming a framework based on a special, multi-tunnel based, predictive, seamless handover solution (Thesis III.1 and III.2). On the other hand I have further extended Host Identity Procotol (HIP) and my already introduced μHIP scheme by developing a HIP-based NEMO protocol (Thesis III.3) for the Host Identity Layer.

4.1. Predictive Handover Management for Multihomed NEMO configurations in IPv6

4.1.1 Overview of predictive mobility management schemes

Multihoming is an advantageous method to provide always-on connectivity in a wireless environment. Mobile clients continuously change their position, which could yield access network failures or connection drops. Mobility management in heterogeneous access architectures is aware of handling the mobility related procedures. IPv6 has built-in support for terminal and network mobility, but these basic solutions do not tackle the problem when handovers provide serious communication outages due to the large number of L1/L2/L3 duties. First, the mobile terminal/router has to find and connect to the new network at L1 and L2 (PHY and MAC), and only after the successful L1/L2 connection it could launch the necessary L3 procedures to obtain the new IPv6 address(es) (with stateless [114] or stateful

autoconfiguration [115]). After the new IPv6 address is set, the binding procedure starts: it binds (registers) its address(es) in the Home Agent, which provides global accessibility. These procedures could easily result in several seconds of handover delay.

Various proposals have been published to shrink the delay caused by handovers. In the next section, I propose to use location information coming from e.g., Global Navigation Satellite System (GNSS), or Geographical Positioning System (GPS) data to speed-up the handover process of multihomed NEMO architectures in heterogeneous access environments. My proposed method is only usable when the mobile terminal or mobile network moves on nearly the same path every time. Public transportation vehicles (trains, trams, buses, trolleys, etc.), cars and trucks traveling on highways/main roads are examples, when this assumption is valid. Random walks in city centers are beyond applicability, and thus our method cannot be applied there.

Using location information for preparing handovers dates back to 2001. In [112] Wang et al. propose to use location information to improve the performance of inter-cell handovers. Their method is limited to L1/L2 handovers, they did not consider IP connectivity. Dutta et al. [113] extended Wang's work recently, also concentrating on the L1/L2 handovers only. Hee-Dong Park et al. [38] proposed first a NEMO scheme which can be used in vehicles travelling on a predetermined route. They store access network information in a database which is used to predict handovers. They have not considered MCoA scenario, though. My method makes use of multihoming as I propose to use MCoA with advanced policy exchange mechanisms. The policy exchange mechanism is based on the recommendations of the IETF's Flow Bindings RFC in Mobile IPv6 and NEMO BS [108], [109]. My solution supports IPv6 only, due to the fact that all related protocols are better implemented in IPv6.

In [116], the authors propose a similar scheme to mine, however this paper lacks the technical description of the system and the handover execution scheme.

4.1.2 GNSS aided predictive handover management for multihomed NEMO configurations

4.1.2.1 General considerations of the proposed scheme

There are two levels of handovers which should be considered independently: L1/L2 handovers, and L3 handovers. L1/L2 level handovers are determined by the access technology currently in use. 3G/HSPA, WLAN, etc. handover delays are due to their respective standard and implementation. However, if the mobile terminal or router contain more than one egress interface, it is possible to use one interface for communication and another one for preparation and execution of L1/L2 handovers. In such a handover scenario sessions should be re-directed between interfaces. Since my solution is based on IPv6, native IPv6 support is a must in all access networks.

L3 level handovers are handled by IP mobility solutions (e.g. MIPv6 or NEMO BS). L3 handovers can be speed-up by launching L3 procedures before L1/L2 handover happens. For this reason I use location information. It could be possible to launch the L3 procedure such that it just finishes as the new network appears. If so, the handover latency becomes lower (down to L1/L2 handover delay) and the service becomes almost ubiquitous. Under perfect circumstances (exploiting overlapping coverage areas and benefits of multi-access devices) the latency can be totally eliminated if the L1/L2/L3 preparations are executed in a predictive and timed manner.

In my IPv6-based NEMO BS extension I propose to use location information coming from e.g., Global Navigation Satellite System (GNSS), or Geographical Positioning System

(GPS) data to speed-up the handover process of multihomed NEMO architectures in heterogeneous access environments.

Thesis III.1. [C19], [C25], [B7] *I have developed a location information aided predictive mobility management framework with an efficient handover execution scheme for multihomed NEMO BS configurations, which combines the benefits of MCoA with a new prediction-driven cross-layer management entity allowing NEMO BS mobile routers to operate using always the best available access networks and to perform seamless handovers when multiple overlapping radio coverages are available.*

The idea behind predictive handover management is very simple: as the node/network moves along a path, it records all access network related data in a database together with the geographical location information. The next time the node/network moves along the same path, based on the geographical information and speed vector, the stored information can be used to predict and prepare handovers before the actual availability of the networks. In the appropriate time, ongoing communication sessions can be seamlessly redirected to some other interface(s) – thus successfully finishing handovers.

The following information should be stored. Network type (WLAN, 3G, WiMAX, etc.), network identifier (e.g. BSSID of WLAN AP, 3G cell identifier, etc.) and IP level information (e.g., network prefix, which can be used to gather the IPv6 address of the node). The first three fields are required: without them it is not possible to prepare handovers in L1/L2/L3 relations. Some additional information, e.g., Signal-to-Noise Ratio (SNR), BandWidth (BW), reliability (how often the network appears at a given geographical location) and Round Trip Time (RTT) are useful for further intelligence and more sophisticated decisions. For instance, the more reliable network should be chosen if several networks are available.

When multiple interfaces are available, MCoA [108] and Flow Bindings [109], [110] solutions can be of use. L3 handover preparation consists of the following components. First of all, a Binding Identification (BID) number is created for each egress interface the Mobile Router (MR) possesses. These BIDs are used as unique identifiers of the interfaces. BIDs are sent in Binding Update (BU) messages to the Home Agent (HA) in order to identify individual bindings of the MR. The HA that receives the BU messages creates a separate binding for each BID (i.e., for each egress interface of the MR). Therefore the MR owns only one Home Address but the bidirectional NEMO tunnels will be distinguishable based on the BIDs. The sole Home Address of the MR requires the introduction of Flow Bindings which directs packet flows to specific egress interface. In the proposed scheme I use Flow Bindings to direct the whole traffic of the MR through one active egress interface. In this way the solution loses the benefits of redundant interfaces, but gains the possibility to use inactive interfaces for handover preparation, i.e., selecting appropriate access network, performing lower layer connections and acquiring new IPv6 addresses.

Therefore the scheme requires several interfaces for operation. Some of the interfaces are used for normal communication (they will be referred as “active”), the others are used for handover preparation (they are termed as “inactive”). The activation of a new interface must be accurately synchronized with the deactivation of the old one. The activation/deactivation procedure means simultaneous reallocation of NEMO BS tunnels. It is performed by properly scheduled flow binding policy control messages on the HA and the MR. The control messages are called Predictive Policy Exchange Messages.

4.1.2.2 The proposed framework and handover execution protocol

The proposed framework (Fig. 23) has three main components: Access Network Prediction (ANP), Handover Manager on the MR (HM-MR) and on the Home Agent (HM-HA). I do not claim all the functional entities are my results; however the overall framework and the design of the predictive handover execution scheme are.

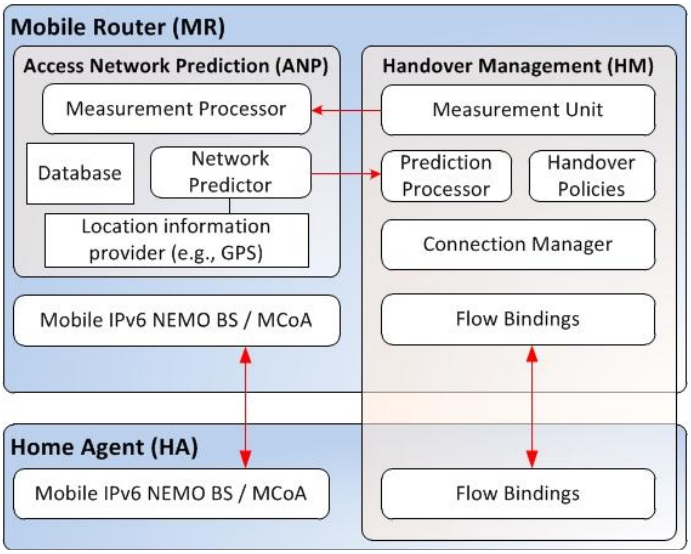


Figure 23: The proposed framework

The left most module on Fig. 23 running inside the Mobile Router is the ANP. The ANP is responsible for 1) maintaining the access network database; 2) sending prediction messages to HM-MR module; 3) reading GNSS information from the GNSS receiver; and 4) processing the network measurement messages received from the HM-MR module. Tasks 3) and 4) are for the maintenance of the access network database which should be continuously extended/updated during the movements of the MR. Based on up-to-date database records and current, precise position/speed information the ANP is able to provide candidate network parameter prediction. The prediction vector is sent to the HM-MR module in an XML message, and the HM-MR measurement messages are also transmitted in an XML format. In order to avoid the explosion of the size of the access network database, the received GNSS coordinates are rounded in the following way: the longitude and latitude values are multiplied by 10,000 and rounded to the closest integer. therefore instead of a continuous space they form a limited set with members called raster points inside a raster net, which plays an important role in the prediction precision (see later).

The Handover Management (HM) module can be divided into two parts depending on which node hosts it. The HM-MR runs on the Mobile Router (Fig. 23) and is responsible for two main tasks. On the one hand HM-MR measures the channel state information and other network parameters of the actually available access networks during the movement of the MR. The scale of the measurable parameters is wide and depends on the decision algorithm to be applied. In my proposal the following parameters are measured, collected and sent periodically to the ANP module for further processing and storing in the database:

- Receive Signal Strength Indicator (RSSI) of UMTS
- Signal/Noise Ratio (SNR) and Basic Service Set Identifier (BSSID) of WLAN
- IPv6 prefix information

On the other hand, HM-MR also prepares predictive handovers by handling MCoA tunnels in a timed manner based on the prediction XML messages received from ANP and the indirect interaction with the NEMO MCoA implementation. In order to achieve this, I proposed a special predictive policy exchange scheme which can inform the Home Agent (i.e., the HM-HA module) about the Mobile Router's intents of future handovers. The periodically received candidate access network predictions supply all the necessary information required for handovers to be initiated by the HM-MR. If a handover event is predicted for the near future (e.g., prediction data reveal that the currently used access coverage will disappear soon), the decision algorithm will choose the destination network and initiate the handover mechanisms. In the proposed framework HM-MR follows a simple rule set to select the designated network from multiple candidates:

- an available WLAN network always has higher priority than 3G/UMTS
- the WLAN with the best SNR value has the highest priority among simultaneously available WLAN networks

The HM-HA module is located on the Home Agent (Fig. 23). The HA itself represents the same functional entity as in the case of standard MIPv6/NEMO/MCoA protocols, but in my scheme it also interacts with the HM-MR module through the HM-HA instance for predictive, timed and flow binding based NEMO MCoA handovers using the Predictive Policy Exchange Messages. The HA is informed about the predicted network prefixes and timing information, and thus changes in flow binding policies can be executed and scheduled before the handover event actually happens.

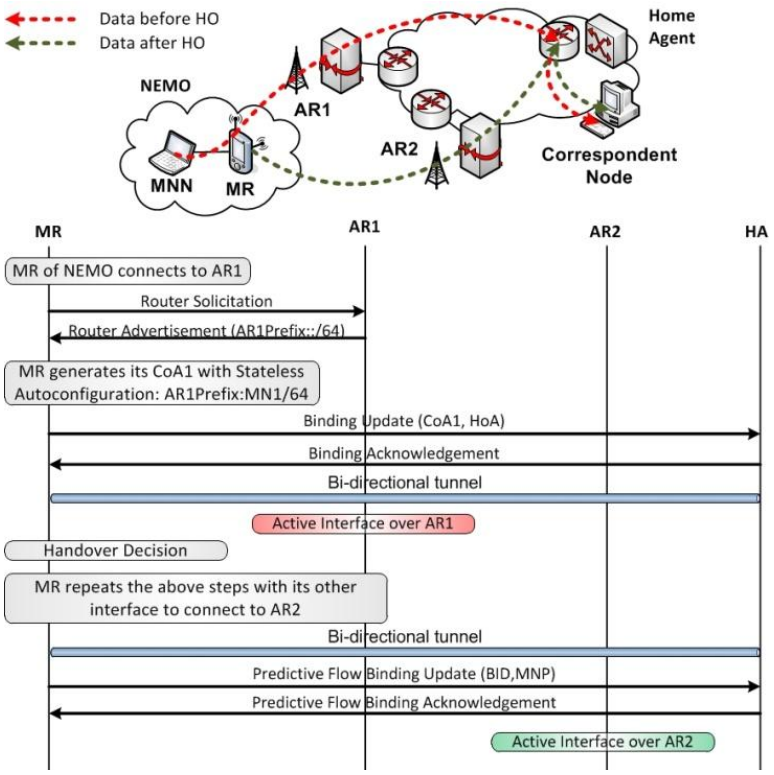


Figure 24: The proposed handover execution protocol

After the decision is made based on the rules defined at the HM-MR module, the

designated network will be chosen and passed over to the Flow Bindings submodule at the MR side. This submodule handles the signaling between the MR and the HA for defining which MR-HA tunnel shall the system switch to and when. It is important to note, that before executing these timed and synchronized policy exchange commands for tunnel/routing adjustments on the MR and the HA entities, the designated network (i.e., a new interface) must be chosen and the L1/L2/L3 preparations must be finished for the selected network. Thanks to the prediction based and multihomed NEMO operation, the MR will be able to finish these preparations (including also the L3 NEMO MCoA tunnel build-up using the binding procedure) before the actual handover event occurs. However, it requires that the candidate networks are overlapping in their coverage.

Based on the GNSS aided predictions the policy exchange commands can be executed at exactly the same time both in the HM-MR and HM-HA modules. It means that all the NEMO traffic will be redirected to the new network defined by the new MCoA tunnel without noticeable packet loss or other QoS disruption. This is only possible because we already have a working Mobile IPv6 connectivity through the new network and all L2/L3 configurations are already performed. The Predictive Policy Exchange message would only carry timed commands to switch the packet flow to a functional, but inactive tunnel. Upon disconnect or failure of the active access network routes and tunnel interfaces are deleted and the next default route with the highest priority is taken to ensure seamless connectivity. Recovering from such failure based on the enforcement of handover policies is out of scope of this document.

The proposed handover execution protocol is detailed in Fig. 24. When the HM decides to perform a handover, in order to use the benefits of MCoA, the following steps are executed. Using one of the inactive interfaces the HM connects to the new access network and establishes a new Mobile IPv6 binding. At this stage, the current and new access networks are both connected and Mobility Tunnels are established between the MR and the HA. Handing over to the new access network is entirely based on Flow Bindings, which in this case means that all flows are moved from one interface to another. To avoid asymmetric routing, the MA and HA has to modify their bindings simultaneously, in a timely manner. The schedule is communicated by the Flow Binding modules in predictive Flow Binding Update/Acknowledgement messages (Fig. 24). When the changes of flow bindings are executed, the new interface is marked as active, while the rest of the communication interfaces are set to inactive mode. The mobile network nodes (MNNs) inside the NEMO will always and transparently use the communication path spanned by the active interface (Fig. 24). Different Handover Policies may have different effects on handover strategies.

4.1.3 Analysis of prediction accuracy in the proposed solution

The proposed framework and handover execution protocol strongly relies on the prediction accuracy which depends on the rasterization scheme working inside the ANP module. That is why I have started to analyze the limitations of the overall architecture inherited by possible wrong positioning on the raster net inside the ANP.

Thesis III.2. [C25], [J16] *I have developed a probabilistic system model for the ANP module and proposed an appropriate rasterization scheme where the probability of wrong positioning on the raster remains below 1%.*

Assume that we have a set of raster points given as $\mathcal{S} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_\infty\}$. $\boldsymbol{x}_i$ represents the ith point which is a geographical position with two coordinates: one on the west-east axis and

one on the north-south axis. $\mathcal{S}$ is an infinite but countable set. The members of the set are constant: they are given by the actual raster size.

Assume that we are at a geographical position $\boldsymbol{x}_0$ ($\boldsymbol{x}_0$can be given by god – no possibility to measure it exactly). We have a GNSS measurement equipment and want to figure out, what $\boldsymbol{x}_0$ is. We make measurements and we get $\boldsymbol{\eta}$ as an estimate, which is not exact of course. $\boldsymbol{\eta}$ is a random number (Gaussian, due to the large number of independent effects), with expectation of $\boldsymbol{x}_0$ and covariance matrix $\boldsymbol{C}$:

$$\mathbb{E}\{\boldsymbol{\eta}\} = \boldsymbol{x}_0 \tag{28}$$

$$Pr\{\boldsymbol{\eta} \leq \boldsymbol{y} | \boldsymbol{x}_0, \boldsymbol{C}\} = \Phi(\boldsymbol{y}, \boldsymbol{x}_0, \boldsymbol{C}) = \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} \frac{1}{\sqrt{2\pi}^2 (det\boldsymbol{C})^{1/2}} e^{(\boldsymbol{z}-\boldsymbol{x_0})^{\mathrm{T}} \boldsymbol{C}^{-1} (\boldsymbol{z}-\boldsymbol{x_0})} dz_2 dz_1 \tag{29}$$

Note that in the last equation we introduced a new notation, Φ. Also note that $\boldsymbol{\eta} \leq \boldsymbol{y}$ means all $\boldsymbol{\eta}$ points where both coordinates are less than or equal to the ones of $\boldsymbol{y}$.

The database uses the raster points only. Thus, based on the measured value $\boldsymbol{\eta}$ we can choose the closest raster point as

$$\boldsymbol{\xi}(t) = \text{argmin}_{\boldsymbol{x} \in \mathcal{S}} \parallel \boldsymbol{\eta}(t) - \boldsymbol{x} \parallel \tag{30}$$

Here, the time dependence have been also added as (t), and $\|.\|$ measures the absolute distance. With the help of God (knowing $\boldsymbol{x}_0(t)$), we would get the perfect estimate $\boldsymbol{x}(t)$ as

$$\boldsymbol{x}(t) = \text{argmin}_{\boldsymbol{x} \in \mathcal{S}} \parallel \boldsymbol{x}_0(t) - \boldsymbol{x} \parallel \tag{31}$$

The first question is the following. What is the probability of making a wrong estimate?

$$Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\} = ? \tag{32}$$

Note that both (30) and (31) are non-linear operations, making it difficult to analyse the problem. The following subsection is about evaluating this probability.

Fig. 25 shows the general geographical setup. As the raster net is self similar, we can put it into the centre of the coordinate system. The area of $\mathcal{T}$ is defined as

$$\mathcal{T} = \{(i,j),\ where\ -a \leq i \leq +a\ \ and\ \ -b \leq j \leq +b\} \tag{33}$$

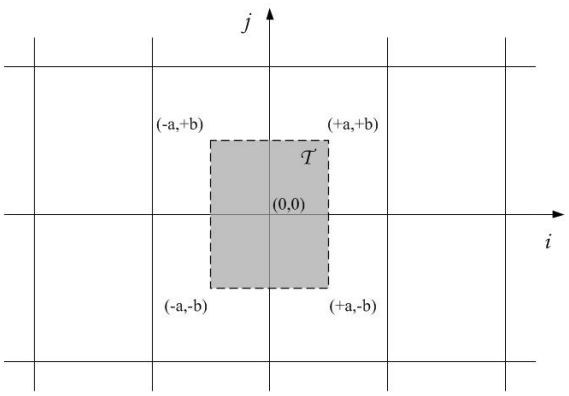


Figure 25: Raster net setup of the probability model

Assuming that the real geographical position ($\boldsymbol{x}_0$) is equally probable at any position, the probability of making a wrong estimate ($Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\}$), equals the probability that the real

position is inside the grey area $\mathcal{T}$ ($\boldsymbol{x}_0(t) \in \mathcal{T}$), and the measured point is outside of $\mathcal{T}$ ($\boldsymbol{\eta}(t) \notin \mathcal{T}$):

$$Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\} = Pr\{\boldsymbol{x}_0(t) \in \mathcal{T} \cap \boldsymbol{\eta}(t) \notin \mathcal{T}\} = 1 - Pr\{\boldsymbol{x}_0(t) \in \mathcal{T} \cap \boldsymbol{\eta}(t) \in \mathcal{T}\} \quad (34)$$

The probability of $\boldsymbol{\eta}$ falling into $\mathcal{T}$ can be computed as

$$Pr\{\boldsymbol{\eta} \in \mathcal{T} | \boldsymbol{x}_0, \boldsymbol{C}\} = \Phi\left(\binom{+a}{+b}, \boldsymbol{x}_0, \boldsymbol{C}\right) + \Phi\left(\binom{-a}{-b}, \boldsymbol{x}_0, \boldsymbol{C}\right) - \Phi\left(\binom{+a}{-b}, \boldsymbol{x}_0, \boldsymbol{C}\right) - \Phi\left(\binom{-a}{+b}, \boldsymbol{x}_0, \boldsymbol{C}\right) \quad (35)$$

Following equation (34), the Bayes' rule and our positioning error constraint, we get

$$0.01 \geq Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\} = 1 - \frac{1}{4ab}\int_{-a}^{+a}\int_{-b}^{+b} Pr\{\boldsymbol{\eta} \in \mathcal{T} | \boldsymbol{x}_0, \boldsymbol{C}\}_{\boldsymbol{x}_0 = \binom{i}{j}} djdi \quad (36)$$

Considering a GPS system for our GNSS measurements with a horizontal positioning error of $\sigma = 5m$ (standard deviation), and taking into consideration that the length in one minute of longitude depends on the latitude (which is ~47.5° N for Budapest), we get that the appropriate raster net is larger or equal to $18.2m$ x $27m$.

4.2. Network Mobility Support in the Host Identity Layer

In order to provide network mobility support in the Host Identity Layer, I have further extended my μHIP framework (already introduced in Thesis I.3) and built a novel, HIP-based NEMO protocol (called HIP-NEMO) upon it.

4.2.1 Overview of novel (not purely IPv6-based) NEMO architectures

Despite the fact that NEMO-BS answers the main questions of network mobility, it still has open issues regarding [117], [118]:

- Optimal routes: By reason of the subservient nation of the HA-MR (parent MR - child MR) liaisons and the encapsulation/tunneling procedures, the routing path in NEMO-BS is highly dependent on the level of nesting resulting suboptimal routes.
- Security: In order to provide security services, NEMO-BS adopts IPSec as its main security convention however it is highlighted that the incorporation of IPSec is insufficient in several scenarios [119].
- Fault-tolerant Home Agents: Since the active connections of a whole moving network are maintained by a single HA entity, the survivability of HAs is one of the critical issues of NEMO-BS.
- Multihoming: A multihomed MNet can benefit the advantages of having redundant links and connectivity gaining fault tolerance, seamless handovers and load sharing. NEMO-BS doesn't include specification for managing multihomed MNets, furthermore the solution is not trivial, especially in more complex nested scenarios.
- Header overhead: Encapsulating packets results in growing header overhead as the level of nesting increases because each packet must be encapsulated (by adding a new IP header) several times.

- Elimination of long delays due to the MR-HA (parent MR - child MR) bidirectional tunnels: The suboptimal routes easily can cause big RTTs thus long packet delays can be observed in several NEMO scenarios.

Many efforts have already been made to provide a more efficient and secure NEMO protocol. Wide variety of different extensions, schemes and methods have been discussed at the IETF NEMO Working Group and at various conferences as well. In this section I give a short overview of the different approaches, in particular the ones which try to present an integrated solution for all the matters and issues of managing moving networks by searching novel, not purely IPv6-based paradigms.

Extensions for NEMO BS solve the problems of multihoming [29], optimal routes and packet overhead [32], [120], and even several security problems [121], but today none of them provides a complete, coherent, integrated framework to address all the issues of moving networks. In order to do this, novel architectures differing from the MIPv6 based NEMO schemes have also been published.

F. Teraoka et al. [122] developed a novel architecture for network mobility management using Location Independent Networking for IPv6 (LIN6). LIN6-NEMO provides network mobility transparency by bringing in a so called Mapping Agent (MA) that manages the location of the MNNs and by performing a procedure in the root MR for overwriting the network prefix of the destination address of packets destined to the MNNs inside a MNet. Despite the fact that vLIN6 enhances the capabilities of the base LIN6-NEMO, several open issues remain open regarding the MA's functionality and the security.

H. Chung-Ming et al. [41] introduced a SIP-based network mobility protocol in order to avoid the problems inherited from MIPv6. The alternative approach of SIP-NEMO extends the SIP framework with three types of new entities called the SIP Home Server, SIP Network Mobility Server and SIP Foreign Server. Based on this new SIP architecture NEMO support can be achieved and even route optimization can be performed between SIP clients. However the solution doesn't handle all the security issues of managing NEMO scenarios.

The first proposal addressing HIP-based NEMO was described in [123]. The authors discuss the basics of a possible NEMO solution that handles HIP aware mobile networks. While this paper is the first work touching mobility of HIP networks as a whole, it does not get into describing the details of the solution.

S. Herborn et al. [124] also presented a HIP-based method for dual layered mobility management controlled by a dynamic context driven heuristics that is responsible for decisions regarding the scheme to be used. This proposal composes not a clean HIP architecture but a hybrid system because the authors assume that mobility will be supported via heterogeneous NEMO architectures created by mobile routers running MIPv6-based NEMO protocols, MNNs with HIP support and CNs also running HIP stack.

In my work I introduced a complete network mobility framework called HIP-NEMO by classifying the mechanisms, defining the algorithms and evaluating the performance of the proposed NEMO solution designed to operate fully in the Host Identity Layer.

4.2.2 HIP-NEMO: Network mobility support in the Host Identity Layer

Although there were some proposals like [123], [124] before HIP-NEMO, my work can be considered as the first complete and pure HIP-based NEMO solution followed by several further extensions and optimizations. In this section introduce my approach enabling a new application of the Host Identity Protocol to provide efficient and secure network mobility

support for HIP-aware moving networks and its terminals even in multihomed and nested scenarios.

Thesis III.3 [C6], [C17], [B5], [J2], [J14], [J17] *I have proposed a Host Identity Protocol based network mobility solution (HIP-NEMO) that introduces a new network entity called Mobile Rendezvous Point (mRVS) in charge of providing NEMO services for HIP-aware MNNs. The scheme eliminates suboptimal user-plane tunneling known in NEMO-BS and ensures efficient communication of the mobile network. I have shown by extensive simulations built on complex protocol models that my proposed HIP-NEMO scheme outperforms the standard NEMO-BS network mobility management solution with an average throughput gain of 211% and provides a significant functional extension to the basic HIP protocol without considerable cost impacts.*

My solution is based on a new network entity called the mobile RVS (mRVS), which holds the role of Mobile Routers and provides certain HIP-based services for nodes in the mobile network. Before I start the detailed description of HIP-NEMO we define some terms that will be used. This is important since I use the NEMO BS terminology in a slightly different meaning. It is necessary to make this differentiation since most of the NEMO related terms were derived from definitions like home network, visited network or Home Agent, which have no real sense to use in a HIP environment.

- *Binding* is a HIP level IDENTIFIER – LOCATOR (i.e., HIT – IP address) couple as opposed to NEMO BS, where binding links two LOCATORs (IP addresses). In my term the *Binding* process is used by the Host Identity Layer to convert HITs to IP addresses and vice versa.
- *Local Fixed Node* (LFN) is a node in the mobile network with a permanent IP address. The node is called *local* as it uses mRVS as its primary Rendezvous Service (RS) provider. In NEMO BS LFN refers to a node, which is in the same home network as the Mobile Router.
- *Local Mobile Node* (LMN) uses mRVS to access the RS but its IP address may change from time to time. In NEMO BS the term refers to a mobile node in the mobile network, whose home network is the same as the one of the mobile router.
- *Visiting Mobile Node* (VMN) uses discrepant RVS as the mRVS to access HIP RS (in NEMO BS VMN is a mobile node in the mobile network, which is assigned to a home link that does not belong to the mobile network).
- *Mobile Network Node* (MNN) is used to refer to all kinds of nodes that may appear in a mobile network (i.e., LFN, LMN, VMN). In NEMO BS terminology MNN means nodes or routers that may appear in a mobile network. In this paper MNNs do not refer to mRVSs.
- *Mobile RVS* (mRVS) is a HIP enabled Mobile Router (MR). All the terms related to MR defined in NEMO terminology are applicable in case of mRVS. Here I mean terms like egress and ingress interfaces, nested mobility terms like root-mRVS (MR), parent-mRVS (MR) and sub-mRVS (MR).
- In NEMO terminology the abbreviation *NEMO* refers to either *NEtwork MObility*, as a networking scenario, and also *MObile NEtwork*, as a networking entity. In this paper I apply NEMO only for the networking scenario. I use MNet to refer to a mobile network, as a networking entity.

Introducing my approach first I highlight the main ideas behind the solution, after which I take a walkthrough over a simple NEMO scenario and explain how HIP-NEMO works in this case. Second, the handover framework of the solution is explained.

4.2.2.1 Protocol architecture of HIP-NEMO

My first consideration was to define the roles and responsibilities of the entity, which will hold the function of mobile router. First, it obviously has to be a HIP-aware node since HIP enabled nodes and Host Identity specific connections have to be managed. Second, since all MNNs are reachable via this entity, it is a rendezvous point (RP) in a HIP aware context. Thus it is a practical choice to have an RVS-like entity in the role of HIP-aware mobile router. I defined a new network entity that meets these requirements, which I call mobile RVS (mRVS) in the rest of the paper.

The mRVS is an RP for MNNs, but not exactly in the same way as standard RVSs. The reason is that mRVS is mobile hence it can provide *mobile rendezvous point* services. This indicates that an mRVS itself must have a normal RVS for employing standard HIP RS. This can be considered as an abstraction of standard HIP RS because in HIP-NEMO I define a RP to another (mobile) RP while in HIP a RP is linked to a concrete HIP node. The mRVS provides the following functions:

- Serves as the primary access point to HIP rendezvous service for LFNs and LMNs.
- Every mRVS acts like a signaling proxy for LFNs and LMNs. The service is offered to nodes directly connected to the particular mRVS, and not for nodes in a nested NEMO. Secure signaling on behalf of other nodes is achieved by HIP-based signaling right delegation [81].
- The standard RVS used by an mRVS is the permanent RP for LFNs and LMNs of the mRVS, which is responsible for registering these nodes at its RVS.
- Finally it communicates with other mRVSs to efficiently handle all the complex NEMO scenarios. On the one hand this communication is based on HIP service discovery [85]. The mRVS frequently sends SAP packets on its ingress interfaces to inform other mRVS about its presence. On the other hand it is possible to define a special inter-mRVS registration mechanism for optimization of nested scenarios.

LFNs and LMNs can operate according to standard HIP in an mRVS-driven NEMO except that they have to be able to delegate their signaling rights to the mRVS. However, this shall be considered as a HIP-based service, rather than a modification of the base protocol.

4.2.2.2 Initialization and connection establishment

The basic HIP-NEMO scenario consists of a single mobile network with one mRVS and a LFN as Fig. 26 shows. The figure also depicts the HIP layer binding of the entities with a number indicating the concrete step of the process, which resulted the particular binding.

Since the LFN is permanently connected to the mRVS it is an obvious choice to enable the mRVS to offer rendezvous service for the LFN. This can be a local policy setting (i.e., LFN knows the HIT and IP address of the mRVS) or can be announced with HIP service discovery mechanism (1). Either initiated by a local policy or by the reception of a SAP packet the LFN initiates the rendezvous registration mechanism with the mRVS. During the process the LFN also delegates its signaling rights to the mRVS. The main purpose of this registration is to get the mRVS to open a new HIT-IP binding entry in its database. Unlike standard RVS entities, which store single {HIT – IP address} mapping records, the mRVS links this information with another IP address. This is a globally routable and topologically correct address, which is allocated and assigned by mRVS to the particular LFN. The role of this address is to provide

global locator for LFN. As Fig. 26 depicts, mRVS assigned IP.LFN.P0 for LFN, while its actual IP address is IP.LFN.P1. One possible way to allocate a proper IP address for LFN is as follows. At registration LFN communicates on IP.LFN.P1. The mRVS simply changes the prefix (P1) of this address to P0, and leaves the remaining part of the address unchanged. Consider IP.LFN.P0 and IP.LFN.P1 as global unicast addresses, introduced in [125]. These addresses consist of two 64 bit long part, as Fig 27 shows.

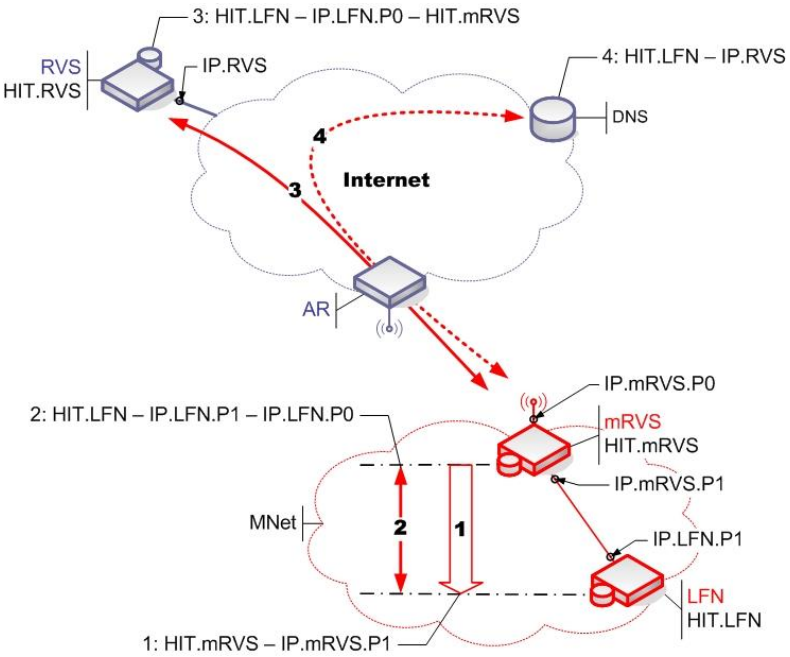


Figure 26: Initialization of a single HIP-NEMO scenario

The address assigned by mRVS to LFN should differ from the real address of LFN in the *Global Routing Prefix* (i.e.P1 and P0). The *Subnet ID* and the *Interface ID* remains the same.

As described above, LFN delegated its signaling rights to mRVS. The first advantage of this process is that it enables mRVS to establish global reachability for the LFN. As Step 3 indicates, the mRVS registers the LFN at its standard RVS using the HIT of the LFN and the IP address assigned it. Moreover, the stored information is indexed by the HIT of mRVS. The role of this index is to enable mRVS to update bindings in RVS with sending only one UPDATE packet. The details are discussed in the next subsection. Finally a DNS record has to be stored that links HIT.LFN to IP.RVS. This is also done by the mRVS and enables correspondent nodes to reach LFN (i.e. through the RVS).

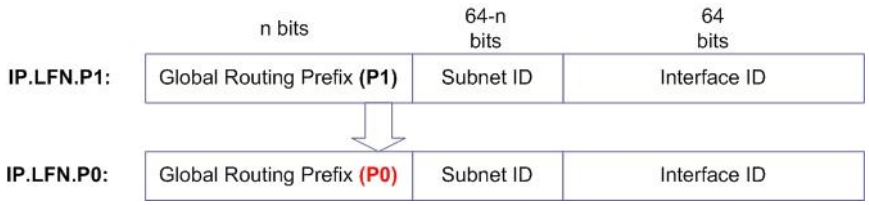


Figure 27: The address allocated by mRVS at LFN registration

Fig. 28 presents the process, in which a CN initiates a HIP association with an LFN. As [23] describes, if a HIP node wants to contact another one, it can initiate the connection in two different ways. If CN is aware of the IP address of its peer, CN can send I1 (i.e., the first

packet of the Base Exchange) directly to this address. In case of unknown peer address CN should send the packet to the serving RVS of the peer.

Let's assume that in the scenario shown in Fig. 28 the CN initiates the connection through the RVS. As described above, the mRVS stored a {HIT.LFN – IP.RVS} record in the DNS thus, after a DNS query, I1 is sent to RVS. The server has the actual HIT – IP binding regarding LFN, according to which the RVS forwards I1 to the mobile network. As I1 reaches the mRVS it changes the destination address prefix (P0) to the actual prefix (P1) of LFN. Before forwarding I1 to LFN the mRVS learns the {HIT.CN – IP.CN} binding. This entry is indexed with the HIT of the LFN, which is used by the mRVS to perform necessary signaling functions (i.e., at sending updates). Finally LFN gets I1, learns the {HIT.CN – IP.CN} binding and continues the Base Exchange by sending R1, which source address is changed from IP.LFN.P1 to IP.LFN.P0 in the mRVS. The R1 packet is routed directly to CN and at reception it obtains the {HIT.LFN – IP.LFN.P0} binding.

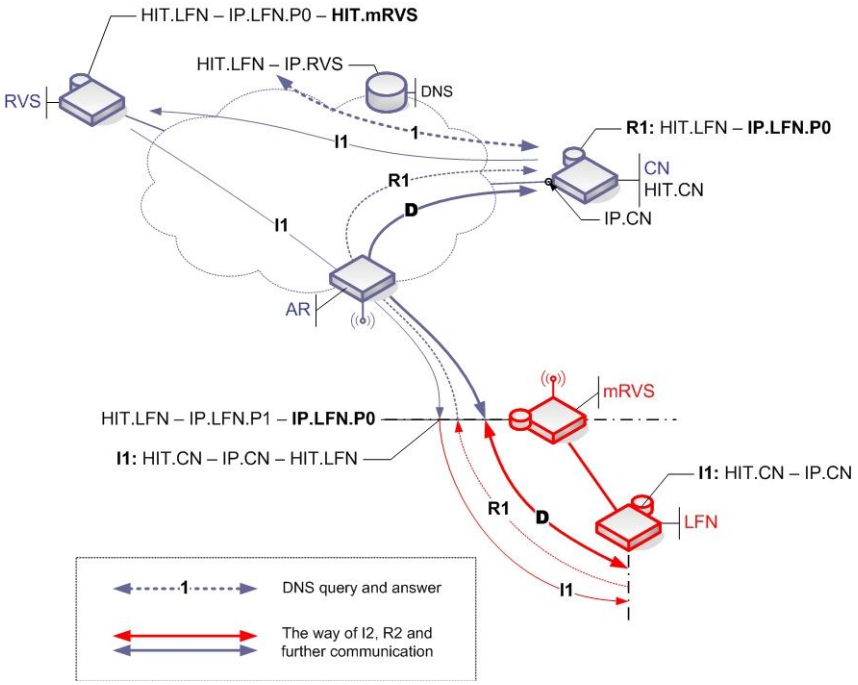


Figure 28: Connection establishment

Finally the Base Exchange is finished and data is ready to transmit. Note that all the packets are routed between LFN and CN on the optimal path but all of them must be processed by the prefix changing function in mRVS. This provides seamless network mobility support for LFNs and ensures optimal routes. However, it has its prize, namely all HIT – IP address binding created at communication partners and at the RVS has to be updated if the whole mobile network changes its point of attachment. The mRVS is in charge to send these updates. On the one hand mRVS has the right to do this (i.e., signaling delegation) and also ensures mobile network handovers transparent to LFNs. On the other hand, if large networks with lots of LFNs are in motion the mRVS might need considerable computational strength to fulfill all update needs. Note that here we mean large but not nested NEMOs because an mRVS is only responsible for nodes directly connected to it. For nodes in a nested subnet the mRVS of this subnet is responsible.

Concluding the protocol description introduced until now I feel worth to highlight the followings. There is a new network entity, mRVS, which is in charge to support seamless NEMO service for LFNs. The mRVS assigns a unique globally routable and topologically correct IP address different for each LFN This address appears in the source address field of

the IP header of all packets traveling out of the mobile network. In the other direction a reversed change of the destination address is accomplished. Finally, mRVS stores information (HIT.CN – IP.CN) about LFNs' communication peers to be able to send the updates on behalf of the LFNs. This might be considered as a HIP level binding list entry.

For LMNs the same process is also applicable, and provides seamless NEMO management for them as well. In case of VMNs my proposed solution cannot provide transparency in terms that every VMN has to inform the network about its presence and vice versa. Furthermore, VMNs have to delegate their signaling rights to the mRVS. A VMN might have active HIP connections and is already registered at an RVS. When it enters in the NEMO it receives SAP packets announcing the presence of the mRVS. In this case the VMN has to make a special registration at the mRVS, which is used to delegate signaling rights and to let mRVS to assign an IP address to the given VMN like it did in case of LFNs or LMNs. The information on CNs of a VMN shall be derived from the UPDATE packets that VMN sends to its CNs, when entering the NEMO. After this, VMNs can be handled in the same way as LMNs. Note that mRVS offers rendezvous service for VMNs (i.e., SAP), but they won't accept it since these nodes have been registered at another RVS before entering the NEMO.

4.2.2.3 Handover framework

Consider a situation when a single mobile network with one mRVS and a LFN connected to it moves away and connects to a new access point (Fig. 29). The egress interface of mRVS gets a new IP address. This also indicates that the mRVS has to assign a new IP address to the LFN. This is necessary since the IP address that the mRVS assigned to LFN needs to be topologically correct unless communication partners won't reach LFN.

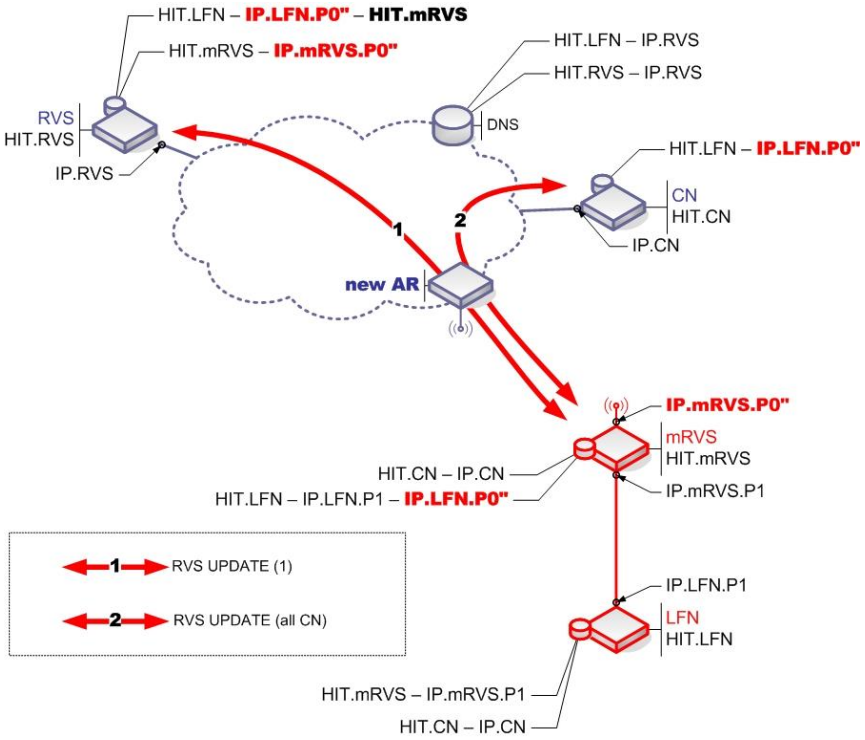


Figure 29: Handover scenario in HIP-NEMO

As mRVS has right to signal on behalf of LFN, it can send HIP UPDATE packets to the RVS and to all of the communication partners of LFN. The UPDATE packet sent to the RVS contains a LOCATOR parameter, which holds only the new prefix (P0") that mRVS uses on its new location. On reception of the UPDATE packet, the RVS will update the binding of the

mRVS. Furthermore, all other bindings are updated that are indexed by the HIT of the mRVS. Note that only the prefix of the addresses in these bindings must be changed. This is applicable since only the prefix field is changed in the new IP address assigned by mRVS to LFNs, compared to the old address. The last 64 bits (assuming /64 networks) remains the same. Note that the "real" IP address of LFN is not changed during the process and thus LFN is not included in the signaling mechanism. The same process can be used in case of LMNs existing in the mobile network.

4.2.2.4 Simulation environment and evaluation results

In order to evaluate my proposed HIP-NEMO scheme and compare its performance against the standard and widespread NEMO-BS solution as reference, I have extended the already introduced HIPSim++ framework [C17] with the novel protocol models and the required special NEMO scenarios.

I have defined one mobile network with one MNN and one mRVS/MR for both the HIP-NEMO/NEMO-BS scenarios, respectively (Fig. 30). This mobile network was able to perform handovers between different access points in the home and foreign networks. In the case of NEMO BS there was a Home Agent (HA) in the topology for the MR, and in the case of HIP-NEMO there was an RVS defined. This RVS acts as the standard RVS for all HIP nodes in the network including the mRVS and the HIP MNN as well. After the simulation start, a CN initiated UDP or TCP connections with the MNN, and then started to send data packets. By using a special node in the topology between the home network (containing the HA and the RVS) and the foreign networks, additional delay was introduced into the test setup.

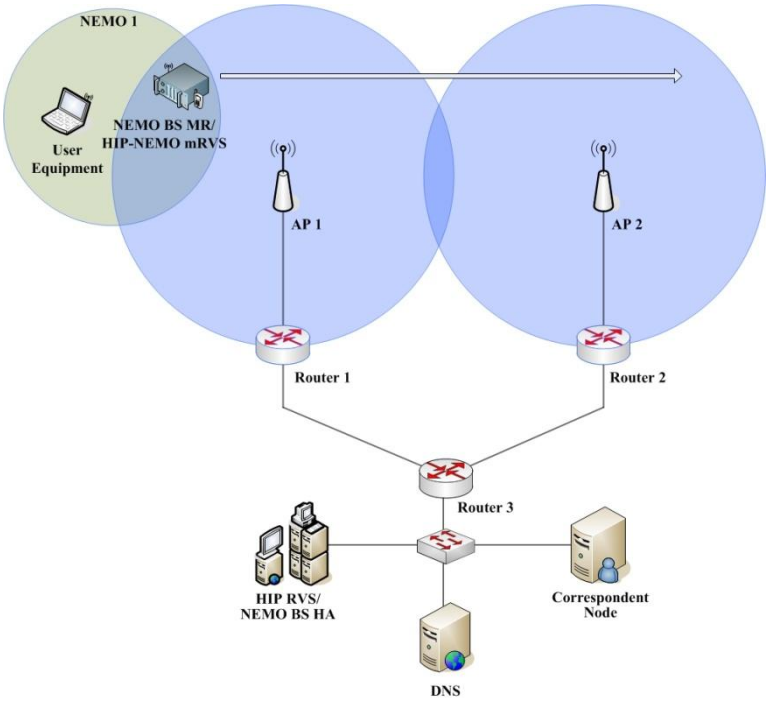
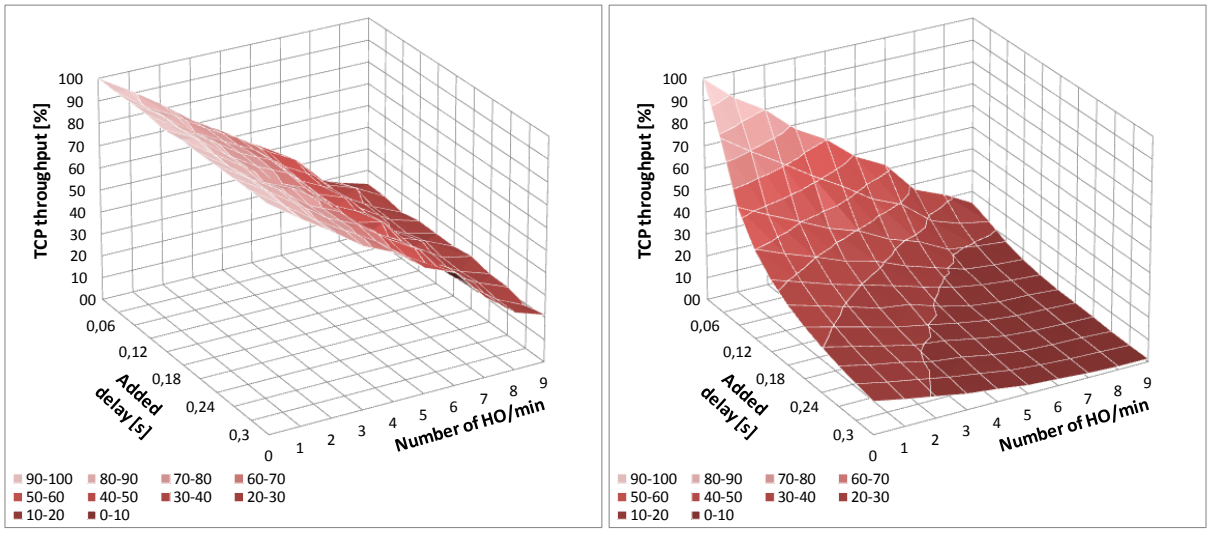


Figure 30: Simulation scenarios for standard NEMO BS mobility and my HIP-NEMO scheme

During the simulation runs I have increased this additional delay from 0 to 300 ms between the HA/RVS and the MR/mRVS respectively, and measured UDP packet loss of one handover in function of the additional delay and different offered data rates (Fig. 32), and TCP throughput for one minute communication in function of the additional delay and different number of foreign-to-foreign network handovers (Fig. 31). The gathered and

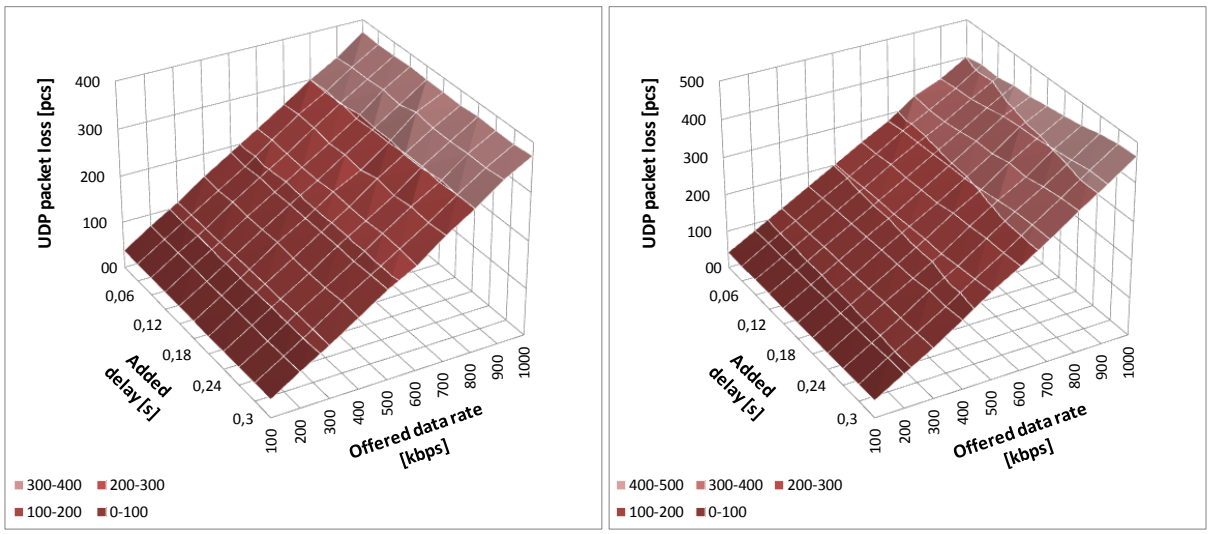
depicted surfaces were rendered using the averages of a total of 10000 independent measurements each.



a) HIP-NEMO c) NEMO-BS

Figure 31: Simulation results for the TCP throughput measurement

TCP results (Fig. 31) show that the frequency of handovers has tremendous impact on both solutions. The TCP throughput during the movement decreases significantly with higher handover frequency. However, as in case of NEMO BS all signalling and data traffic flows through the HA, the rate of that decrease is higher than it is when HIP-NEMO is used. My solution builds near-optimal routes between communicating entities and eliminates user-plane tunnels from the communication path which also results in the fact that user-plane performance will not be tainted by the increasing topological distance between the home and foreign networks (i.e., between the RVS and the HIP supported mobile network). The cumulated average throughput gain of HIP-NEMO is around 211%.



a) HIP-NEMO c) NEMO-BS

Figure 32: Simulation results for the UDP packet loss measurement

UDP results (Fig. 32) further strengthen the above statements. The effects of one handover are independent from the added delay in case of HIP-NEMO, while NEMO-BS

shows notable degradation when the mobile network is more far away from home. The average performance gain of HIP-NEMO in terms of lost UDP packets for the high data rate and high added delay cases is more than 19%.

During my evaluation I was focusing on the network mobility management efficiency of HIP-NEMO, as it is its primary advantage by using direct routes between MNNs and their CNs. Furthermore, my proposal scales well with large and complex mobile networks. The main advantage of HIP-NEMO is that it presents network mobility management integration to HIP by extending the usability of the base protocol. On the other hand HIP-NEMO benefits from being derived from HIP. Namely the effective security and mobility-multihoming framework of the base protocol is inherited and applied efficiently.

In the current Internet where hosts are identified according to their IP addresses, the true advantage we get from HIP is a strong identification based on the cryptographical Host Identities. HIP enabled hosts can prove their identity by owning the private key part of their asymmetric Host Identity and signing data with it. With cryptographic identities, HIP enables authentication between end-points. Initialization of a HIP association is designed to protect the responder from Denial of Service (DoS) attacks. Communication confidentiality with HIP is established by encrypting the payload data.

Currently, the specified encryption format is ESP. Furthermore, HIP protects the integrity and confidentiality of payload data as well as integrity of control packets. HIP control packets can also be used to carry cryptographic certificates. Certificates can be used for authentication or authorization purposes by the peer host or intermediate entities. The latter property is a key issue, when considering secure signaling right delegation.

MNNs delegate their signaling rights to one (or more i.e., multihoming) mRVS in a secure way by sending registration packets that hold the correspondent certificate. Basic HIP security functions and secure delegation of signaling rights together provide secure location update.

Since signaling rights are delegated in a secure way and base HIP signaling messages are signed by the sender, location update signaling is protected. Service discovery that is used by mRVSs to advertise their presence and the rendezvous service they are offering for MNNs shall be considered as the security bottleneck of the solution. When a HIP host or a mRVS of a nested subnet receives a SAP packet from the network, either as a result of an active service discovery, or passively, it cannot know if the service provider is trustworthy or not. The SDP packet is unprotected, which makes it vulnerable. An attacker can modify the packet, or an attacker can send the packet using someone else's IP address and HIT. However, there are situations when nodes or moving networks have no other choice but to trust other nodes because there are no other means for them to connect to the Internet. Note that MNNs delegate their signaling rights to the mRVS directly. In a singlehomed environment this is the only way for MNNs to connect to the public network. On the other hand, the decision of to whom shall I delegate my right of signaling becomes a more complex problem in multihomed environments.

Chapter 5
Schemes for Distributed and Flat Mobility Management

Mobile Internet has recently started to become a reality for both users and operators thanks to the success of novel, extremely practical smartphones, portable computers with easy-to-use 3G USB modems, and attractive business models. More and more people are willing to access Internet related services while being on the move, requesting seamless, ubiquitous connection anytime and anywhere. Based on the current trends in telecommunications, vendors prognosticate that overall mobile data traffic is expected to grow nearly 11 fold between 2013 and 2018 [1]. It is also expected, that the increase of mobile Internet traffic will be higher compared to the fixed Internet traffic in the forthcoming years, most dramatically due to new entrant, data-hungry mobile entertainment services like mobile TV, video and music [2], and new application types, such as M2M (machine-to-machine) communications including e-health services, vehicle communications, remote control and monitoring services [4], and P2P file and multimedia sharing. In order to accommodate the future Internet to the anticipated traffic demands, technologies applied in the radio access and core networks must become scalable to advanced future use-cases.

There are many existing solutions aiming to handle the capacity problems of current mobile Internet architectures caused by the mobile traffic data evolution. Reserving additional spectrum resources is the most straightforward approach for increasing the throughput of the radio access, and also spectrum efficiency can be enhanced thanks to new wireless techniques (e.g., High Speed Packet Access (HSPA) [126], and Long Term Evolution (LTE) [127]). Heterogeneous systems providing densification and offload of the macro-cellular network throughout pico, femtocells and relays [128] or WiFi/WiMAX [129] interfaces also extend the radio range. However, the deployment of novel technologies providing higher radio throughput (i.e., higher possible traffic rates) easily generates new usages and the traffic increase may still accelerate. As because today's mobile Internet architectures have been originally created for voice services and later extended to support packet switched services only in a very centralized manner, the management of this ever growing traffic demand is not a simple task to deal with. The challenge is even harder if we consider fixed/mobile convergent architectures managing mobile customers by balancing user traffic between the widest scale of different access networks.

The growing number of mobile users, the increasing traffic volume, the complexity of mobility scenarios, and the development of new and innovative IP-based applications require network architectures able to deliver all kind of traffic demands seamlessly assuring high quality of service. However, the strongly centralized nature of current and planned mobile Internet standards (e.g., the ones maintained by the IETF or by the collaboration of 3GPP) prevents cost effective system scaling for the novel traffic demands.

Motivated by the above reasoning, novel mobile architectures and accompanied protocols started to emerge where bottlenecks from packet communication is removed by eliminating user-plane anchors from the network and bringing pure IP routing close to the mobile terminals in terms of physical location in the architecture [16]. Decentralized, robust, self-configuring and self-optimizing network structures are envisioned with reduced operation expenditure (OPEX), improved system capacity and energy efficiency. A discussion of the disadvantages of anchor-based and advantages of end-to-end mobility protocols regarding the architecture together with the most important scalability problems can be found in [17], [43]. To enhance scalability of the core network in mobile Internet architectures, the Ultra Flat architecture (UFA) has been introduced by Khadija Daoud et al. [17]. Their proposal reduces the number of network nodes to only one serving node called the UFA Gateway (UFA GW)

and traditional user and control plane functions are distributed in such UFA GWs deployed at the edge of the architecture, close to the subscribers. The main characteristic of this proposal is that the execution of handovers is managed by the network via the Session Initiation Protocol (SIP) [44] operating within the frame of the IP Multimedia Subsystem (IMS) [130], [C20], [J5]. Even though SIP is a very powerful signalling solution for UFA, it is not applicable for non-SIP (i.e., legacy Internet) applications and the published SIP-based UFA scheme also does not comply with ITU-T's recommendation of requirements for ID/Loc separation in future networks that allows the network layer to change locators or even protocols without troubling upper layer communication sessions [18].

This motivated me to work on an alternative signalling scheme for the Ultra Flat Architecture based on the promising ID/Loc separation method of the Host Identity Protocol. My contribution to the developed HIP-based Ultra Flat Architecture (UFA-HIP) lies in the design of its novel, HIP mobility management, signalling delegation, context transfer and cross-layer interworking based system framework (Thesis IV.1), in the development of a proactive, distributed, 802.21 MIH / HIP-based handover preparation and execution protocol (Thesis IV.2), and in the performance evaluation of the UFA-HIP mobility management scheme by extensive simulations (Thesis IV.3).

5.1. HIP-based Ultra Flat Architecture (UFA-HIP)

5.1.1 Traffic Evolution Characteristics and Scalability Problems of the Mobile Internet

The continuous growth of mobile broadband traffic volume is inevitable. Furthermore, the evolving technologies applied in access networks, user terminals, and user applications will seriously affect the traffic patterns as they are common today. These symptoms form our motivation to analyze the driving forces behind the trends, and to present the scalability problems of mobile Internet caused by them.

One of the most important reasons of the traffic volume increase in mobile telecommunications is demographical. According to the current courses, world’s population is growing at a rate of 1.2 % annually, and the total population is expected to be 7.6 billion in year 2020. This trend also implies a net addition of 77 million new inhabitants per year [3]. Today, over 25% of the global population – this means about two billion people – are using the Internet. Over 60% of the global population – now we are talking about five billion people – are subscribers of some mobile communication service [2]. Additionally, the number of wireless broadband subscriptions is about to exceed the total amount of fixed broadband subscriptions and this development becomes even more significant considering that the volume of fixed broadband subscriptions is gathering much slower.

The number of mobile handhelds in use with broadband subscriptions will increase drastically in the near future, but due to the fixed Internet connection replacement still the mobile broadband connected notebooks and laptops equipped with USB or integrated wireless modems will realize the most significant part of the mobile data traffic. Wired Internet applications and services are subjects of transition to wireless and mobile broadband networks since Internet customers expect to have comparable user experience and level of comfort on the move, as they were at home or in the office. This is a fundamental driver for mobile broadband penetration and appearance of new device types.

The expansion of wireless broadband subscribers not only inflates the volume of mobile traffic directly, but also facilitates the growth in broadband wireless enabled terminals. However more and more devices enable mobile access to the Internet, only a part of users is attracted or open to pay for the wireless Internet services meaning that voice communication

will remain the dominant mobile application also in the future. Despite this and the assumption of [3] implying that the increase in the number of people potentially using mobile Internet services will likely saturate after 2015 in industrialized countries, the mobile Internet subscription growth potential will be kept high globally by two main factors. On the one hand the growth of subscribers continues unbrokenly in the developing markets: mobile broadband access through basic handhelds will be the only access to the Internet for many people in Asia/Pacific. On the other hand access device, application and service evolution is also expected to sustain the capability of subscriber growth.

The most prominent effect of services and application evolution is the increase of video traffic: it is foreseen that due to the development of data-hungry entertainment services like television/radio broadcasting and VoD, 69.1% of mobile traffic will be video by 2017 [1]. A significant amount of this data volume will be produced by mobile Web-browsing which is expected to become the biggest source of mobile video traffic (e.g., YouTube). Cisco also forecasts that the total volume of cloud applications and services such as Netflix, YouTube, Pandora, and Spotify will reach almost 90 percent of all consumer traffic (fixed and mobile) by the year 2018, producing a substantial increase of the overall mobile traffic 13-fold between 2012 and 2017 [131]. Video traffic is also anticipated to grow so drastically in the forthcoming years that it could overstep Peer-to-Peer (P2P) traffic. Emerging web technologies (such as HTML5), the increasing video quality requirements (HDTV, 3D, SHV) and special application areas (virtual reality experience sharing and gaming) will further boost this process and set new challenges to mobile networks. Since video and related entertainment services seems to become dominant in terms bandwidth usage, special optimization mechanisms focusing on content delivery will also appear in the near future. The supposed evolution of Content Delivery Networking (CDN) and smart data caching technologies might have further impact on the traffic characteristics and obviously on mobile architectures.

Another important segment of mobile application and service evolution is social networking. As devices, networks and modes of communications evolve, users will choose from a growing scale of services to communicate (e.g., e-mail, Instant Messaging, blogging, micro-blogging, VoIP and video transmissions, etc.). These services are getting more and more widespread and as they often use a mix of voice, video and text transmission, they generate considerable traffic. In the future, social networking might evolve even further, like to cover broader areas of personal communication in a more integrated way, or to put online gaming on the next level deeply impregnated with social networking and virtual reality. Despite the fact that social networking applications and services are envisioned to produce much more data sessions compared to pure video services, they will not produce more traffic since they are not so bandwidth-consumptive. The data traffic volume of mobile voice services will also become less significant compared to other data-hungry communication modes but in terms of gross profits, voice services and basic texting (e.g., SMS) will remain crucial for the operators.

Even though video seems to be a major force behind the current traffic growth of the mobile Internet, there is another emerging form of communications called M2M (Machine-to-Machine) which has the potential to become the leading traffic contributor in the future. M2M sessions accommodate end-to-end communicating devices without human intervention for remote controlling, monitoring and measuring, road safety, security/identity checking, video surveillance, etc. Predictions state that there will be 152.2 million cellular M2M devices by 2016 with little traffic per node but resulting significant growth in total, mostly in uplink direction [132]. The huge number of sessions with tiny packets creates a big challenge for the operators. Central network functions may not be as scalable as needed by the increasing number of sessions in the packet-switched domain.

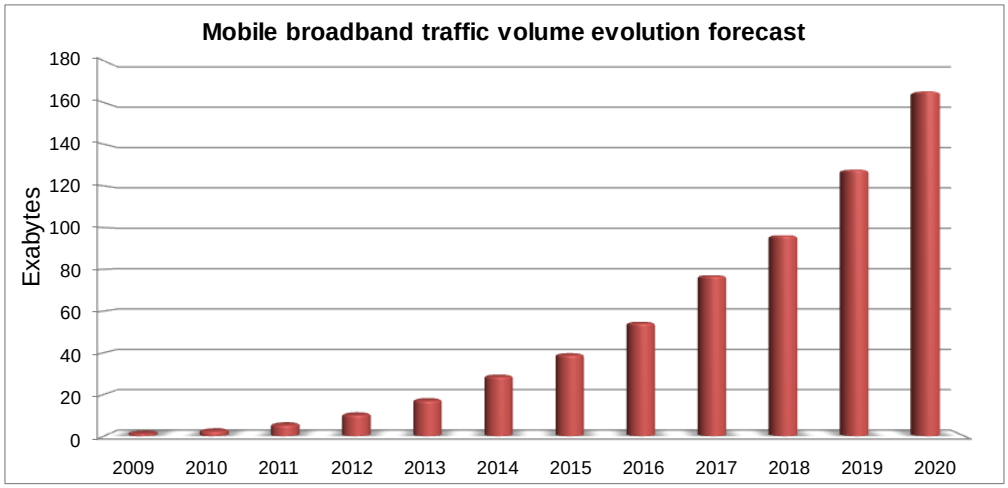


Figure 33: Mobile Traffic Forecast [J9]

As a summary I can state that the inevitable mobile traffic evolution is foreseen (Figure 33) thanks to the following main factors: growth of the mobile subscriptions, evolution of mobile networks, devices, applications and services, and significant device increase potential resulted by the tremendous number of novel subscriptions for Machine-to-Machine communications.

However, existing wireless telecommunication infrastructures are not prepared to handle this traffic increase, current mobile Internet was not designed with such requirements in mind: mobile architectures under standardization (e.g., 3GPP, 3GPP2, WiMAX Forum) follow a centralized approach which cannot scale well to the changing traffic conditions.

On the one hand user plane scalability issues are foreseen for anchor-based mobile Internet architectures, where mechanisms of IP address allocation and tunnel establishment for end devices are managed by high level network elements, called anchor points (GGSN in 3GPP UMTS, PDN GW in SAE, and CSN for WiMAX networks). Each anchor point maintains special units of information called contexts, containing binding identity, tunnel identifier, required QoS, etc. on a per mobile node basis. These contexts are continuously updated and used to filter and route user traffic by the anchor point(s) towards the end terminals and vice versa. However, network elements (hence anchor points too) are limited in terms of simultaneous active contexts. Therefore in case of traffic increase new equipments should be installed or existing ones should be upgraded with more capacity.

On the other hand, scalability issues are also foreseen on the control plane. The well established approach of separating service layer and access layer provides easy service convergence in current mobile Internet architectures but introduces additional complexity regarding session establishment procedures. Since service and access network levels are decomposed, special schemes have been introduced (e.g., Policy and Charging Control architecture by 3GPP) to achieve interaction between the two levels during session establishment, modification and release routines. PCC and similar schemes ensure that the bearer established on the access network uses the resources corresponding to the session negotiated at the service level and allowed by the operator policy and user subscription. Due to the number of standardized interfaces (e.g., towards IP Multimedia Subsystem for delivering IP multimedia services), the interoperability between the service and the access layer can easily cause scalability and QoS issues even in the control plane.

As a consequence, architectural changes are required for dealing with the ongoing traffic evolution: future mobile networks must specify architecture optimized to maximize the end-user experience, minimize CAPEX/OPEX, energy efficiency, network performance, and to ensure mobile networks sustainability.

5.1.2 The UFA-HIP System Framework

As I introduced above, it is highly expected that due to their centralized (anchor based) design, mobile architectures currently being under deployment or standardization would not scale particularly well to efficiently handle all the challenges. It is also anticipated that mobility management tasks of advanced scenarios cannot be tackled effectively if IP address will continue to remain both locator (for packet routing) and identifier (for referring to a host or session): the semantically overloaded nature of the Internet Protocol must be obviated by identifier/locator (ID/Loc) separation [18], [80]. A novel mobile architecture should be created focusing on two main goals. On the one hand scalability issues in packet communication must be tackled by removing user-plane anchors. On the other hand service establishment, security and mobility procedures must be optimized by distributing them from centralized nodes and by introducing ID/Loc separation. However the above mentioned reconstruction and optimization of current architectures seems to be inevitable, it cannot be implemented without strict attention to the compatibility with legacy applications and services, introducing a wide variety of new performance and functional constraints.

The basics of such a redesigned mobile architecture were firstly defined by Khadija Daoud et al. in [17], [43], [133] under the name of Ultra Flat architecture (UFA). UFA represents the ultimate step towards flattening IP-based core networks, e.g., the Evolved Packet Core (EPC) in 3GPP. The objective of the UFA design is to distribute core functions into single nodes at the edge of the network, e.g., the base stations. Certain control functions could remain in the core, e.g., to support 3GPP and IMS roaming or to centralize the subscriber information base. The intelligent nodes at the edge of the network are called UFA gateways. Daoud et al. focus on the handover procedure problems of UFA [17]. If the first IP router is located in the access network, mobility introduces frequent IP level handovers, especially in dense areas. The authors have developed a Session Initialization Protocol (SIP) based handover procedure for UFA. It has been proven by analysis [17] and in a testbed [43] that seamless handovers can be guaranteed for SIP-based applications. SIP Back-to-Back User Agents (B2BUAs) in UFA GWs can prepare for fast handovers by communicating the necessary contexts, e.g., the new IP address before physical handover. This scheme supports both mobile node-(MN) and network-decided handovers.

The session establishment procedure and the integration of IMS and UFA have already been investigated [133]. In 3GPP, IMS [130], [C20], [J5] facilitates service and network convergence by separating the service level from the access layer. This introduces a two-level session establishment procedure. First, the MN and the correspondent node (CN) negotiate the session parameters on the service level, then Policy and Charging Control ensures that the bearer established in the access layer uses the resources corresponding to the negotiated session. The problem is that the service level is not directly notified about access layer resource problems, and e.g., it is difficult to adapt different components of the same service to the available resources in the access layer. In UFA, access layer resource information is present in the close neighborhood of UFA GWs. UFA GWs (B2BUAs) can influence the negotiated parameters during the SIP session establishment and update. Consequently, in addition to enhancing scalability, purely SIP based UFA is entirely controlled by the operator, and integrates Quality of Service (QoS) in its establishment and mobility procedures.

Interworking with Internet applications is a major requirement for mobile operators. In converged networks that use SIP control, an important problem is that many applications preferred by users apply other protocols for session establishment, e.g., Hypertext Transfer Protocol (HTTP). I refer to those applications as non-SIP applications. Therefore, in this section I propose an alternative system framework for the Ultra Flat Architecture based on the promising ID/Loc separation method called Host Identity Protocol (HIP), driving by the idea of realizing IP mobility management in a layer lower than the application layer.

Thesis IV.1. [C21], [C22], [J6], [J9], [J12], [J13] *I have proposed a Host Identity Protocol based system framework for the Ultra Flat Architecture (called UFA-HIP), which completely eliminates centralized IP anchors between Point of Access (PoA) nodes and correspondent nodes, places network functions at the edge of the transit and access networks (close to PoAs) in the UFA-HIP GWs, integrates 802.21 MIH and HIP features to provide efficient inter-GW mobility management, and incorporates signalling delegation and context transfer to reduce the number of HIP Base Exchanges (BEX) between the MN and the network and within the network and also to remove overhead from wireless links by shifting significant part of signalling overhead of MNs from the air interface to the wired UFA-HIP segments.*

The proposed HIP-based Ultra Flat Architecture (UFA-HIP) system framework defines seven main building blocks (Fig. 34): 1) several access networks (both wired and wireless), 2) an IP/MPLS transit network, 3) HIP capable UFA GWs (UFA-HIP GWs) controlling main network functions, 4) an optimized terminal attachment scheme with cross-layer access authorization, 5) a session establishment protocol, 6) a handover initiation, preparation, decision and execution subsystem based on the IEEE 802.21 Media Independent Handover (MIH) standard [134] and extended HIP functionalities, and 7) a HIP-based control network.

Heterogeneous access networks provide the air interface for MNs making them able to connect to the core infrastructure (and to the Internet) anytime, anywhere. Besides to support IEEE 802.21 mechanisms there are no other restrictions regarding the access technologies to be used in this framework: any kind of access system can be applied in any kind of heterogeneous setup. The IP/MPLS transit network is the operator's backbone including routers and core network elements (for service and configuration provision, 802.21 services etc.), and natively connecting UFA to the global backbone (i.e., to the Internet).

In this system centralized IP anchors between PoAs and correspondent nodes are totally removed, and network functions are placed at the edge of the transit and access networks (close to the PoAs) in the UFA-HIP GWs. The solution applies HIP for IPsec security association (SA) establishment between the MNs and UFA-HIP GWs, and between the UFA-HIP GWs. IP-level handovers are prepared and executed using HIP delegation services [C21] and CXTP-based context transfer [135]. The main tasks of UFA-HIP GWs:

1. Performing fast cross layer (L2 and HIP-level) access authorization.
2. Actively interacting with hosts through delegation-based HIP and IPsec association management and context transfer for optimized message exchange in HIP-based UFA mobility and multihoming operations. (Note that the proposed system framework transports end-to-end flows between MNs and CNs in a hop-by-hop manner. The middle-hops are the UFA-HIP GWs, i.e., the delegates of the end peers).
3. Performing the actual mapping/routing between outer header IPsec tunnels based on inner header identifiers.
4. Coordination of resource allocation, load balancing, and handover decisions with the help of the UFA-HIP cross-layer module (MIHU in the MIH taxonomy).

I proposed the use of HITs in inner IP headers for the identification of flows in UFA-HIP GWs, with the same purpose as the Control Plane Header (CPH) in [136]. Without delegation, maintaining end-to-end security associations (SAs) between every communicating peers would be required, as in the SPINAT-based frameworks [136]. Note, that there is a trade-off between the delegation-based and SPINAT-based alternatives, i.e., my proposed solution alternative introduces an extra-header in every packet, but reduces signaling at the MNs, the

second requires SPINAT-based middleboxes (i.e., UFA-HIP GWs), and MN-initiated signaling for the maintenance of a high number of HIP and IPsec associations.

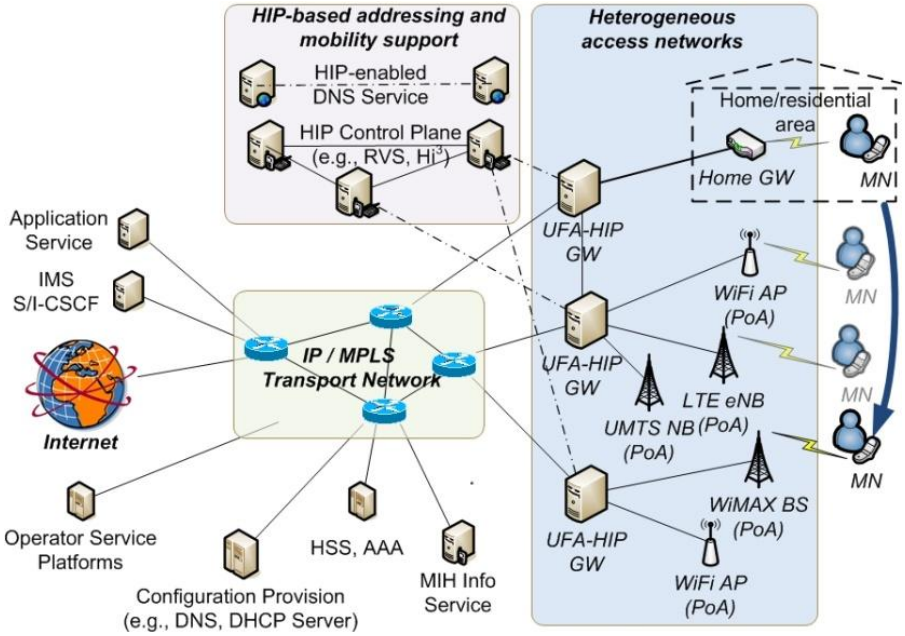


Figure 34: UFA-HIP: The proposed HIP-based Ultra Flat Architecture system framework

There are control functions which are not part of the UFA-HIP GWs and remain in the core network. The optimal location of these functions is subject to further research. Such functions are IP Multimedia Subsystem (IMS), the Home Subscriber Server (HSS), the authentication, authorization and accounting (AAA) servers, service and configuration provision (DHCP), Media Independent Information Service (MIIS). Existing service platforms and application servers remain centralized as well. All core functions and the UFA-HIP GWs are connected with IP networks.

The IEEE 802.21 MIH management subsystem inside my proposed framework handles handover preparation issues and relating signaling tasks in order to initiate proactive HIP handover procedures in the UFA and to support network and mobile controlled handover decision. UFA-HIP GWs are PoS, but often non-PoA entities. According to the MIH standard [134], UFA-HIP GWs must communicate over Reference Point 5 (RP5) with PoAs and over RP3 with MNs. In my system framework, network initiated 802.21 handover preparation procedures are triggered by the serving UFA-HIP GWs (refer to Appendix C.2 in [134]). RP3 and RP5 messages are sent over L3 [137], and protected by HIP and IPsec.

The control network (HIP-based addressing and mobility support) contains a HIP-compatible Domain Name System [84] for resolving domain names to host identities and/or locators depending on the actual situation. In addition there is the HIP Control Plane which stores and distributes dynamic and presumably frequently changing binding information between host identities and locators of all actively communicating (mobile) hosts in UFA-HIP. This control plane might be a conventional RVS park or a complete distributed HIP signalling architecture like Hi^3 [138]. The records managed here are provided by the UFA-HIP GWs using their own global locators as location information to be bounded with identities of their actively interacting partners. The control of the above functions brings cross-layer HIP modules in the UFA-HIP GWs, MNs and CNs.

HIP BEX and Update procedures deal with dynamic negotiation of IPsec security associations between the MN and the UFA-HIP GW to protect user data and mutually authenticate the MN and the network. The handover preparation and initiation subsystem

handles handover preparation issues and relating signalling tasks in order to initiate proactive HIP handover procedures in the UFA and to support both network and mobile controlled handover decisions. The handover execution procedure is started by the source UFA GW (S UFA GW). HIP and IPsec contexts are established between the target UFA GW (T UFA GW) and the MN's CNs, furthermore, between the target UFA GW and the MN, using the mediation of the S UFA GW. This is possible due to the delegation of HIP signalling rights from the MN and from the target UFA GW to the source UFA GW. Context Transfer Protocol [135] is used to transfer the HIP and IPsec contexts from the source UFA GW to the target UFA GW and the MN. As the contexts are in their place the MN is notified by the handover preparation and initiation subsystem to attach to the new PoA.

5.2. Distributed Handover Management Protocol for UFA-HIP

5.2.1 Overview of Distributed Mobility Management

Flat mobile networks not only require novel architectural design paradigms, special network nodes and proprietary elements with peculiar functions, but also require certain, distinctive mobility management schemes sufficiently adapted to the flat way of operation by being distributed in nature. In fact such distributed mobility management mechanisms and the relating methods form the key routines of the future mobile Internet designs. The importance of this research area is also emphasized by the creation of a new IETF non-working group called Distributed Mobility Management (DMM) in August 2010, aiming to extend current IP mobility solutions for flat network architectures.

However, current mobility management solutions rely on centralized architectures employing anchor nodes for mobility signaling and user traffic forwarding. In 3G UMTS architectures centralized mobility anchor is implemented by the GGSN nodes that handle traffic forwarding tasks using the apparatus of GPRS Tunneling Protocol (GTP). The similar centralization is noticeable in Mobile IPv6 [6] where the Home Agent –an anchor node for both signaling and user plane traffic– administers mobile terminals' location information (i.e., the bindings between temporary and persistent IP addresses), and tunnels user traffic towards the mobile's current locations and vice versa. Several enhancements and extensions such as Fast Handovers for Mobile IPv6 (FMIP) [111], Multiple Care-of Addresses Registration [108], Network Mobility (NEMO) Basic Support [15], Dual-Stack Mobile IPv6 [139] were proposed to optimize the performance and broaden the capabilities of Mobile IP, but all of them preserve the centralized and anchoring nature of the original scheme. Micromobility and localized mobility solutions like Hierarchical Mobile IPv6 [9], Proxy Mobile IPv6 (PMIP) [10] or ABMF (Section 2.1.2) try to tackle the problem, but cannot completely overcome to architectural boundaries. There are also alternate schemes in the literature aiming to integrate IP-based mobility protocols into cellular architectures and to effectively manage heterogeneous networks with special mobility scenarios. Cellular IP [11] introduces a gateway router dealing with local mobility management while also supporting a number of handoff techniques and paging. A similar approach is the handoff-aware wireless access Internet infrastructure (HAWAII) [48], which is a separate routing protocol to handle micromobility. Terminal Independent Mobility for IP [140] combines some advantages from Cellular IP and HAWAII, where terminals with legacy IP stacks have the same degree of mobility as terminals with mobility-aware IP stacks. Authors of [141] present a framework that integrates 802.21 Media Independent Handover [134] and Mobile IP for network driven mobility. However, these proposals are also based on centralized functions and generally rely on MIP or similar anchoring schemes.

Some of the above solutions are already standardized [142]–[144] for 3G and beyond 3G architectures where the architectural evolution is in progress: E-UTRAN (Evolved Universal Terrestrial Radio Access Network) or LTE (Long Term Evolution) base stations (eNodeBs) became distributed in a flatter scheme allowing almost complete distribution of radio and handover control mechanisms together with direct logical interfaces for inter-eNodeB communications. Here, traffic forwarding between neighboring eNodeBs is temporarily allowed during handover events; however, traffic anchoring operations remain centralized thanks to e.g., S-GW, PDN-GW, Local Mobility Anchor and Home Agent, responsible for maintaining and switching centralized, hierarchical and overlapping system of tunnels towards mobile nodes. Also, offloading with Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO) extensions [145] cannot completely solve this issue: mobility management mechanisms in current wireless and mobile networks anchor the user traffic relatively far from users' location. This eventuates the implementation of packet encapsulations over the whole network, needs centralized mobility anchors to keep mapping information up-to-date, and results in centralized, unscalable data plane and control plane with non-optimal routes, overhead and end-to-end packet delay even in case of motionless users, centralized context maintenance, single point of failures, and in deployment issues of cache and content servers for Content Delivery Networks (CDN). To solve these entire problems and questions novel, distributed and dynamic mobility management approaches may be envisaged, applicable either to intra-technology or to inter technology mobility cases.

The basic idea is that anchor nodes and mobility management functions of wireless and mobile systems could be distributed to multiple locations in different network segments, hence mobile nodes in any of these locations could be served by a close entity.

First, core-level distribution is achievable, meaning that mobility anchors are topologically distributed covering specific geographical area but still remain in the core network. A good example for this is the Global HA to HA protocol [146], which extends MIP and NEMO in order to remove their link layer dependencies on the Home Link and distribute the Home Agents in Layer 3, at the scale of the Internet. DIMA (Distributed IP Mobility Approach) [147] can also be considered as a core-level scheme by allowing the distribution of MIP Home Agent (the normally isolated central server) to many and less powerful interworking servers called Mobility Agents (MA). These new nodes have the combined functionality of a MIP Home Agent and HMIP/PMIP Mobility Anchor Points. The administration of the system of distributed MAs is done via a distributed Home Agent overlay table structure based on a Distributed Hash Table (DHT) [149], which creates a virtual Home Agent cluster with distributed binding cache that maps a mobile node's permanent identifier to its temporary identifier.

Second, mobility functions and anchors could be distributed in the access part of the network. For example in case of pico- and femto cellular access schemes it could be very effective to introduce Layer 3 capability in access nodes to handle IP mobility management in that way and to provide higher level intervention and even cross-layer optimization mechanisms. The concept of UMTS BSR [16] realizes such an access-level mobility management distribution scheme where a special network element called BSR (Base Station Router) is used to build flat cellular systems. BSR merges the GGSN, SGSN, RNC and NodeB entities into a single element: while a common UMTS network is built from a plethora of network nodes and is maintained in a hierarchical and centralized fashion, the BSR integrates all radio access and core functions. Furthermore, the BSR can be considered a special wireless edge router that bridges between mobile/wireless and IP communication. In order to achieve this, mobility support in the BSR is handled at three layers: RF channel mobility, Layer 2 anchor mobility, and Layer 3 IP mobility. The idea of Liu Yu et al. [150] is quite similar to the BSR concept. Here a node called Access Gateway (AGW) is introduced to

implement distributed mobility management functionalities in the access level. The whole flat architecture consists of two kinds of elements, AGW on the access network side and terminals on the user side. Core network nodes are mainly simple IP routers. The scheme applies DHT and Loc/ID separation: each mobile node has a unique identifier (ID) keeping persistent, and an IP address based locator (Loc) changed by every single mobility event. The (Loc,ID) pair of each mobile is stored inside AGW nodes and organized in a DHT manner. When a mobile node moves to a new network (i.e., under a new AGW) it changes its IP according to the new network and updates the (Loc,ID) pair in the DHT to make sure that the mobile node owning a particular ID can be reached by looking up the IP in the DHT of (Loc,ID) pairs.

Third (and last) type of DMM application scenarios is the so-called host-level or peer-to-peer distributed mobility management where once the correspondent node is found, communicating peers can directly exchange IP packets. In order to find the correspondent node, a special information server is required in the network, which can also be centralized or distributed. A good example for host-level schemes in the IP layer is MIPv6 which is able to bypass the user plane anchor (i.e., Home Agent) based on route optimization mechanisms like [151], such providing a host-to-host communication method. End-to-end mobility management protocols working in higher layers of the TCP/IP stack such as TCP-Migrate [152], Stream Control Transmission Protocol (SCTP) [153], Session Initiation Protocol (SIP) [44] or Host Identity Protocol (HIP) [20] can also be efficiently employed in such schemes.

5.2.2 802.21 MIH and HIP-based handover initiation, preparation, execution and completion

Today's mobility management protocols (e.g., Mobile IP, NEMO BS and Proxy Mobile IP without route optimization) do not separate signaling and user planes which means that all control and data packets traverse the centralized mobility anchor. As because the volume of user plane traffic is much higher compared to the signaling traffic, the separation of signaling and user planes together with the distribution of the user plane but without eliminating signaling anchors can still result in effective and scalable mobility management. This is exploited the proposed UFA-HIP framework where a relatively simple inter-UFA GW protocol can be used thanks to the centralized HIP signaling plane, but the user plane is still fully distributed. Mobile IP based DMM solutions also rely on the advantages of this partial distribution concept when they implement route optimization and such separate control packets from data messages after a short period of finishing the route optimization procedure.

In order to develop an appropriate mobility management protocol for UFA-HIP I extended the base protocol to support UFA GW centric signaling delegation based handover operation, and also to integrate IEEE 802.21 MIH with Host Identity Protocol. The IEEE 802.21 MIH [134] standard specifies a unified framework for proactive handover control in heterogeneous architectures (i.e., 802.3, 802.11, 802.16, 3G networks). It supports event and command service (ES, CS) mainly used for local and remote link-layer event monitoring, and information service (IS) collecting static information on access networks. The previous services enable network and MN-controlled handover decisions, i.e., target L2 Point of Access (PoA) selection. The standard defines procedures for PoA resource availability checks, resource reservation, and release. The handover execution protocols and decision algorithms are outside the scope of the standard. Point of Services (PoS) are network elements that communicate directly with the MN, and can assist in handover decision. In my proposed solution, UFA GWs are PoS, but often non-PoA entities (i.e., no Layer 2 link is available between the MN and the non-PoA UFA GW).

Thesis IV.2. [C21], [C22], [J6], [J9] *I have designed a proactive, distributed, 802.21 MIH and HIP-based handover initiation, preparation, execution and completion protocol for UFA-*

HIP. The proposed technology generally supports flat architectures, minimizes end-to-end path length for user traffic, and keeps the mobility signalling load in the backhaul and core segments.

The proposed distributed mobility management scheme anticipates live registration of MN to the network (including live registration to the serving UFA-HIP GW's signaling delegation services), and live registration between target and source UFA-GWs. The handover protocol starts with the handover initiation.

During the 802.21 MIH handover initiation phase (illustrated in Figure 35), the UFA mobility management algorithm decides to initiate the handover process to one of the candidate UFA GWs. Within this phase, the source UFA GW configures the serving access interface of the multimode terminal with the set of QoS parameters required for the serving access link, using MIH procedures. As a result, the serving access interface periodically notifies the registered MIH user (i.e., the UFA-HIP cross-layer module in the source UFA GW) about its QoS parameters. Based on this information, the special algorithm inside the UFA-HIP module has sufficient information about the serving access network and, if necessary, can trigger the handover preparation phase before connectivity is lost.

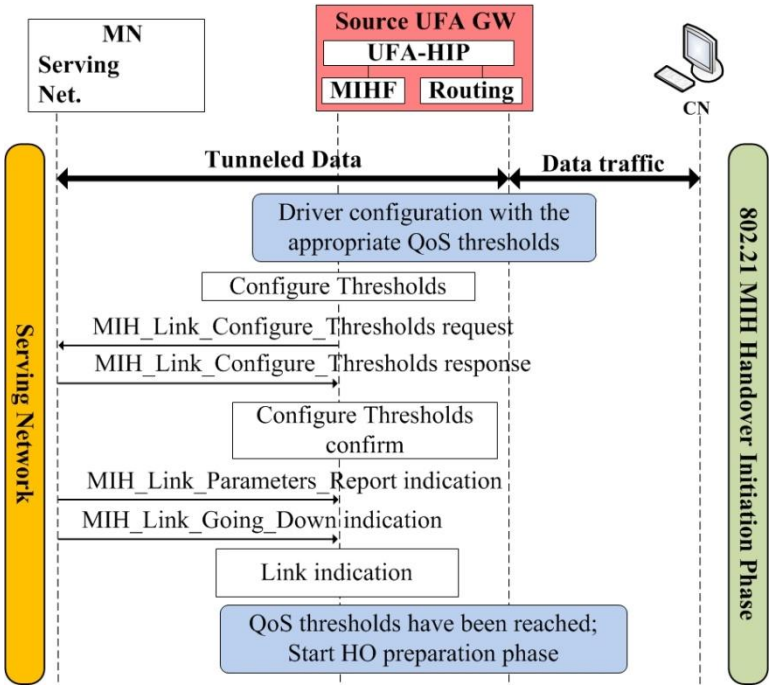


Figure 35: 802.21 MIH handover initiation phase for UFA-HIP

After receiving this trigger message starts the 802.21 MIH handover preparation phase (Fig. 36) with the following sub-phases.

1. Discovery: during this phase, the list of candidate UFA GWs is obtained through the 802.21 Media Independent Information Service (MIIS) [134], which collects information about the candidate access networks, such as their identifiers, L2 addresses, accounting information, etc. UFA GWs may also maintain a local MIIS database.
2. Query: in this phase, the mobility decision algorithm acquires all QoS metrics for all available candidate UFA GWs.
3. Selection: the mobility decision algorithm running either on the network or in the terminal side, decides for the target network (i.e., T UFA GW).

Fig. 36 illustrates the 802.21 MIH handover preparation phase for a network initiated handover. The S UFA GW queries the MIIS about the available neighbouring networks, then asks the MN to narrow the list of candidate access networks, and finally checks the available resources at each C UFA GW. Thereafter, it decides the selected T UFA GW for the handover procedure.

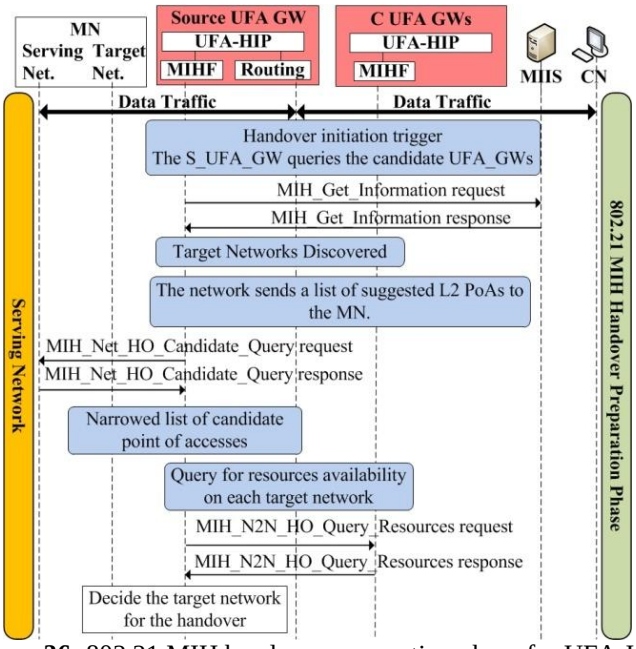


Figure 36: 802.21 MIH handover preparation phase for UFA-HIP

After the selection of the target PoA and the T UFA GW starts the HIP-base handover preparation. First, the necessary HIP and IPsec contexts are proactively established in the network by the S UFA GW using Type 1 and Type 2 HIP delegation services [C21] (please see the explanation of these services and their signaling in Table 1).

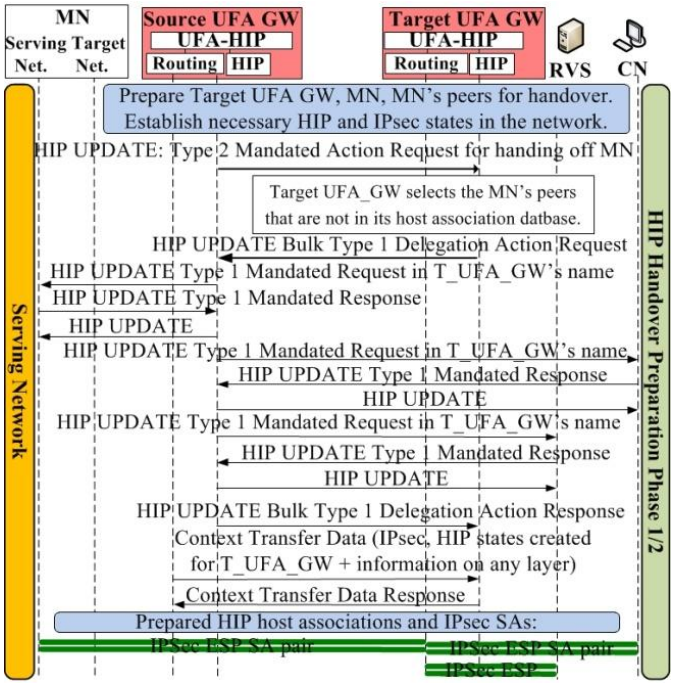


Figure 37: HIP-based handover preparation phase 1/2 for UFA-HIP

The prerequisites of these procedures are that the T UFA GW must register to the Type 1 Delegation service of the S UFA GW, in order to delegate HIP and IPsec association establishment. Furthermore, the S UFA GW (or the MN) must subscribe to the Type 2 Delegation service of the T UFA GW, to authorize the T UFA GW to update the MN's location at the MN's active peers, i.e., its CNs or the UFA GWs of its CNs and the RVS.

As depicted in Figure 37, the S UFA GW initiates a Type 2 Mandated Action Request on behalf of the MN for handing off MN's sessions. It triggers a bulk Type 1 Delegation Action Request sent back to the S UFA GW. With this request the T UFA GW authorizes the S UFA GW for the establishment of HIP and IPsec connections with the MN's peers in the name of the T UFA GW. Then the S UFA GW sends the security contexts to the T UFA GW using CXTP protocol protected with IPsec. Hence, the number of HIP BEX procedures can be reduced and replaced by HIP Updates. After the successful context transfer, the T UFA GW updates the traffic forwarding policies for the MN at the CNs, RVS, and the MN, as illustrated in Fig. 38.

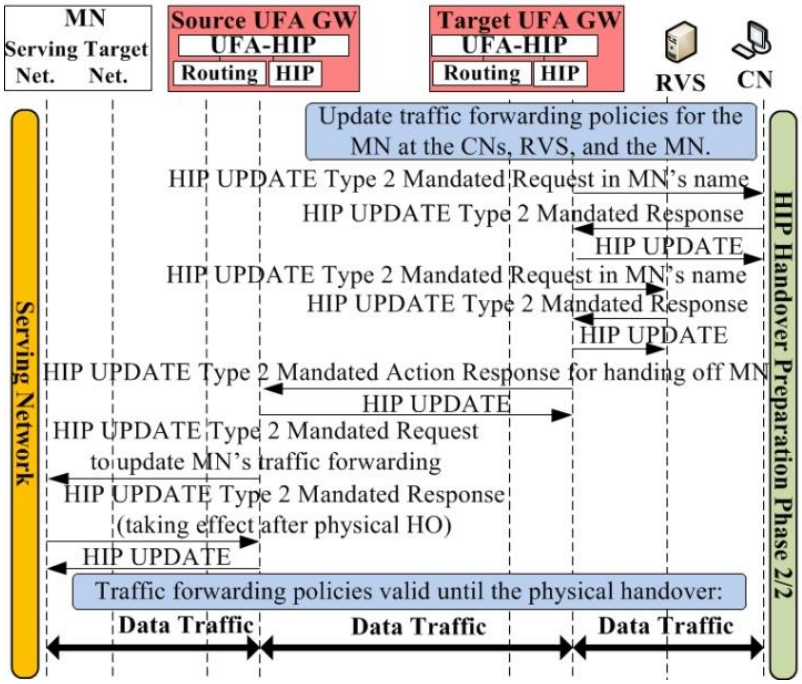


Figure 38: HIP handover preparation phase 2/2 for UFA-HIP

Firstly, Type 2 Mandated Action Requests are sent by the T UFA GW to the CNs and the RVS in the MN's name. After updating the MN's peers, the T UFA GW informs the S UFA GW with a Type 2 Mandated Action Response to prepare for the redirection of the sessions. The T UFA GW updates its HIT-based traffic forwarding table [J6] to receive traffic from MN's peers and send packets towards the S UFA GW. The S UFA GW also updates the MN's and its own local HIT-based traffic mapping table: the traffic coming from the MN, related to the sessions that will be handed off soon, must be mapped to the IPsec tunnel that has the T UFA GW on the other side. The MN delays the activation of forwarding its traffic to the T UFA GW. Therefore, the traffic of the MN passes through the source and target UFA GW until the physical handover completes.

Figure 39 illustrates the handover execution and completion phase for the proposed scheme. After HIP handover preparation phase, L2 handover execution procedure is initiated by the MIH_N2N_HO_Commit and MIH_Net_HO_Commit request messages of 802.21 MIH [134], towards the T UFA GW and the MN, respectively. Then, the MN attaches on L2

to the target L2 PoA. This procedure could contain fast L2 re-authentication schemes, e.g., ERP [154]. The last phase is initiated by the MN when it physically attaches to the target L2 PoA and UFA GW. The MN signals to the T UFA GW that the handover was successfully executed and S UFA GW and L2 PoA can release the resources maintained for the MN's handed off sessions.

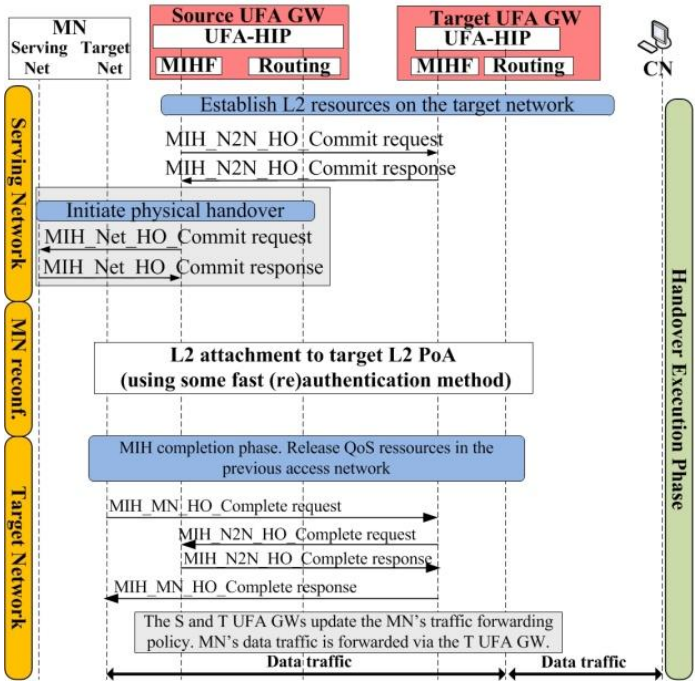


Figure 39: Handover execution and completion phase for UFA-HIP

This last phase is executed by the 802.21 MIH protocol's MIH_MN_HO_complete procedure: after the reception of the MIH_MN_HO_complete request message the T UFA GW requests the S UFA GW to release the resources maintained for the MN's handed off sessions by sending a MIH_N2N_HO_complete request. Finally, the traffic forwarding policy must be updated for the MN in the source and target UFA GWs to exclude the S UFA GW from the path.

Table 1: Explanation of the applied HIP-based Delegation Service messages [C21]

HIP Parameter	Description
Delegation Establishment Request	The Delegator sends to the Delegate for itself or on behalf of another node in order to request Type 1/2 delegation service using HIP REG REQ parameter. Authorization Certificate chain of the acquiring node must be included in HIP NOTIFICATION parameter(s).
Delegation Establishment Response	The Delegate sends to the Delegator in order to acknowledge or reject Type 1/2 delegation service establishment using HIP REG RESP or REG FAILED parameter.
Delegation Action Request	The Delegator sends to the Delegate for itself or on behalf of another node in order to request HIP and/or IPsec association creation or update. In case of Type 1 Delegation Service the state information will be transferred to the Delegator. For Type 2 Delegation Service, the states resulted by the action will be created and further maintained by the Delegate.
Delegation Action Response	The Delegate sends to the Delegator in order to report the Type 1/2 delegation action results in HIP NOTIFICATION parameter(s).
Mandated Action Request	The Delegate sends to 3rd party node(s). For Type 1 Delegation Service HIP and/or IPsec associations will be created by the Delegate and transferred to the Delegator. In case of Type 2 Delegation Service, new HIP and/or IPsec states are created on behalf of the Delegator by the Delegate and/or traffic mapping rules will be updated. HIP NOTIFICATION parameters are used to transfer the required information such as supported IPsec SPI values of the Delegator, global locator(s) of the Delegator, list of supported HIP and IPsec transforms, traffic mapping rules, Delegator peer list, configuration and service registration parameters, etc.
Mandated Action Response	3rd party node(s) send to the Delegate in order to report Type 1/2 mandated action results in HIP NOTIFICATION parameter(s).
Context Transfer Data (CTD)	Sent by the Delegate to Delegator, and includes feature data (i.e., HIP and IPsec context data).
Context Transfer Data Reply (CTDR)	Sent by Delegator to Delegate, indicating success or failure of context transfer.

5.2.3 Simulation Environment and Evaluation Results

The evaluation of the developed scheme was performed in the extended version of the INET/OMNeT++ based open source HIPSim++ [C17] simulation environment already introduced in Thesis I.3.

Thesis IV.3 [C17], [J14], [J17] *I have shown that the proposed proactive UFA-HIP handover preparation and execution protocol reduces the handover latency with an average 67% and thus almost totally eliminates the effects of frequent inter-GW mobility events in distributed or flat mobile networking architectures. I have also revealed the scale of the benefits exploited from the scheme by UDP and TCP applications. The number of lost UDP packets is 55% less in average, while the average TCP throughput gain of the distributed scheme is above 60% compared to the legacy solutions.*

I have used standard HIP and MIPv6 mobility management solutions as reference: the mobile host (MN) changed its PoA by connecting to another Wi-Fi access point (AP) due to its movement. As the APs were connected to different access routers advertising different IPv6 prefixes, the IPv6 address of the MN was changed after reattachment. In the HIP case, standard RFC 5206 mechanisms were applied to handle this mobility situation by running the HIP UPDATE process [22]. In the MIPv6 case two sub-cases were implemented: MIPv6 with and without routing optimization (RO ON and OFF, respectively), where return routeability [6] is applied / not applied on the top of MIPv6 binding update procedure. During the simulation built-in TCP and UDP application models were used to generate traffic between the MN and its Correspondent Node (CN). I have introduced a special router node providing an average RTT of 300 ms between the MN and the CN / HIP RVS / MIPv6 Home Agent to simulate Internet-wide communication. A simple Domain Name Service model was applied used to simulate DNS procedures, but they were initiated only before connection establishment (i.e., HIP BEX).

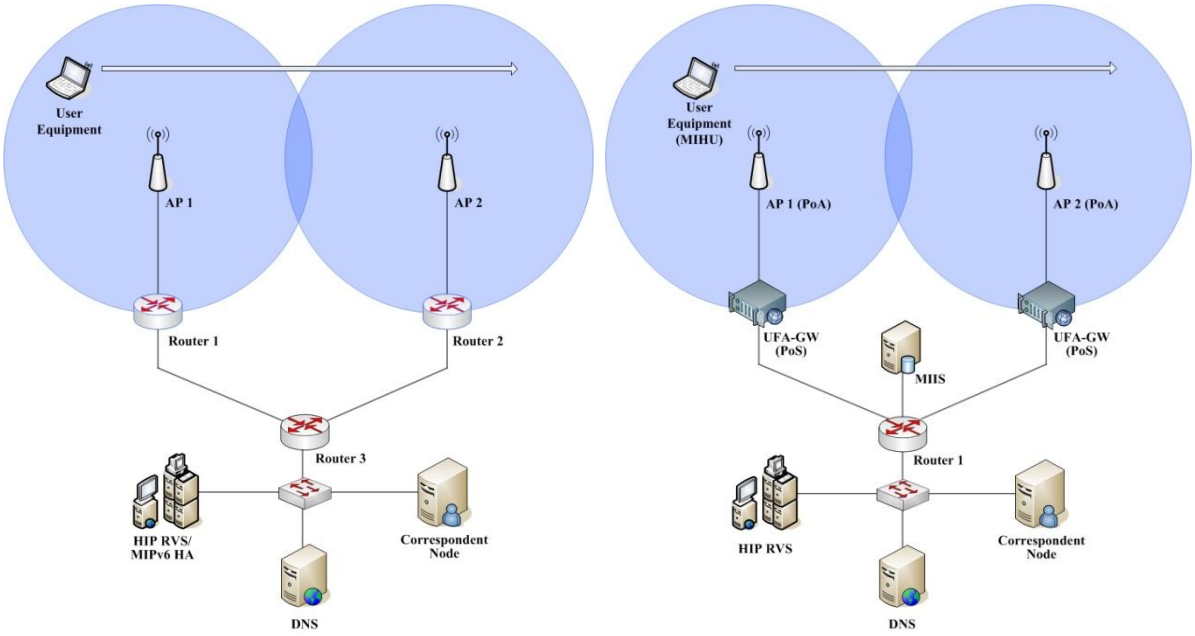
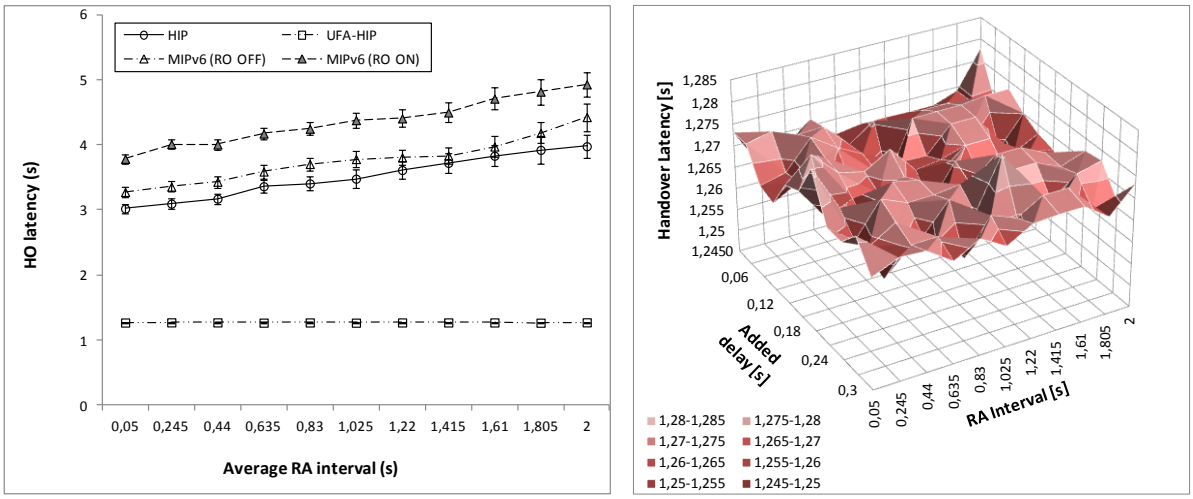


Figure 40: Simulation scenarios for MIPv6/HIP (left) and UFA-HIP (right) schemes

For the UFA-HIP scenario the difference lies in the introduction of UFA-HIP GWs and their advanced signaling delegation based functions: two HIP-capable UFA GW nodes

replace the legacy access routers and control their PoAs (Wi-Fi AP 1 and 2). In this scenario the MN uses active Signal-to-Noise Ratio measurements and a threshold value to trigger handover preparation. HIP functions (signaling delegation and context transfer) were implemented as extensions to HIPSim++, while the model of 802.21 MIH framework was built on the Notification Board toolset of INET/OMNeT++ [87].



a) HO latency vs. RA interval

b) HO latency vs. delay between S and T UFA GW

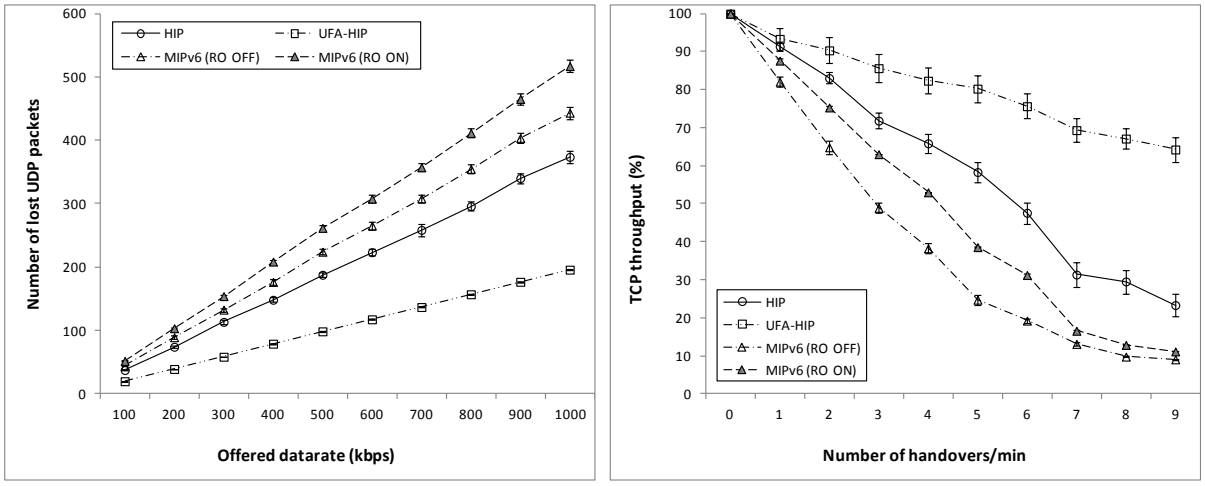
Figure 41: Handover latency of the UFA-HIP scheme

In all the above scenarios the MN is able to migrate between the different APs with a constant speed such provoking handovers situations. By inducing 100 independent handovers during simulation runs I have measured three key performance indicators in three different sub-scenarios. *Sub-scenario A* measures Handover Latency defined here as the time elapsed between loosing the connection at the old AP and completing the handover execution on the MN side while connected to the new AP. Fig. 41/a presents the Handover Latency as the average of the 100 handover series for every RA interval within its 99% confidence interval. Measurements show that UFA-HIP handover performance is independent of the subsidiary IP layer mechanisms (i.e., delays of acquiring IP address, duplicate address detection, etc.) and keeps service interruption delay slightly above 1 sec. It means that the handover latency is caused only by the physical reattachment procedures in UFA-HIP (Wi-Fi AP re-association in our simulations). Measurements show that the service interruption delay of UFA-HIP is independent from the configuration delay in the target network (i.e., RA interval), and about 67% smaller than the reference protocols case in average, thanks to the advanced proactive operation which basically reduces the handover disruption to the Layer 2 (re-)attachment delay. Fig. 41/b presents the Handover Latency in function of the RA interval and the delay between source and target UFA-HIP GWs. The graph depicts that the performance of my proposed proactive handover solution is independent from the topological distance of the S and T UFA GWs.

Fig. 42/a introduces results of sub-scenario B for every evaluated protocol, and shows how much UDP packet was lost during a handover in a HIP, MIPv6 RO ON, MIPv6 RO OFF, and UFA-HIP based system. The points on the graph represent the average UDP packet loss of 100 handovers for every offered datarate value and depicted within a 99% confidence interval. The differences and similarities of the examined protocols' handover performance are clearly observable in the UDP transport layer. The power of my proactive, context-transfer based distributed solution designed for ultra flat architectures is highlighted by the fact that

the number of lost UDP packets is 55% less in average for the UFA-HIP case compared to the legacy HIP and MIPv6 aggregate performances.

Fig. 42/b depicts the TCP throughput proportion of the four protocols under analysis in a one minute communication session between the MN and the CN experienced at different handover frequencies from 0 to 9. The gain of my UFA-HIP solution is eye-catching especially when the circumstances are deteriorating (i.e., the number of handovers is increasing): in case of the highest handover frequency UFA-HIP shows more than 175% gain in TCP throughput, but also the average gain of the advanced distributed scheme is above 60%.



a) UDP packet loss during one HO b) TCP throughput vs. HO/min

Figure 42: Performance of UDP and TCP applications in the UFA-HIP handover scheme

With the help of extensive simulations I proved that the handover latency and the number of lost packets during handoffs can be significantly reduced while the TCP throughput can be considerably increased in case of my UFA-HIP proposal, meaning that relevant improvements in handover performance can be achieved besides the enhanced scalability when applying intelligent distributed HIP gateways in the mobile network. It is an important result as optimization of handover performance is one of the key problems for flat networks: unlike in hierarchical and centralized architectures which usually provide efficient fast handover mechanisms using Layer 2 methods, in flat architectures IP-based mobility management protocol must be used. Since all the PoAs are connected directly to the IP core network, hiding mobility events from the IP layer is much harder. Another aspect of this challenge is that the most widespread IP mobility management solutions are mainly designed for macromobility environments and induce relatively long handover latency. As disruptions during mobility events is not acceptable for real-time applications like VoIP and live video broadcast, high-performance IP mobility management with advanced micromobility-like extension is required for flat IP networks.

In my UFA-HIP system framework high scalability is achieved because centralized anchors – the main performance bottlenecks – are removed, and traffic is forwarded in a distributed fashion. The flat nature also provides flexibility regarding the evolution of broadband access, e.g., the range extension of RANs with unmanaged micro-, pico- and femtocells, without concerns of capacity in centralized entities covering the actual area in a hierarchical structure.

Failure tolerance/resistance, reliability and redundancy is also refined and strengthened as no such single points of failure exist, and the impact of possible shortfalls of the distributed

network elements (i.e., UFA GWs) can smoothly narrowed to a limited, local area. Also no complex failure recovery operations are required: if a UFA GW in the flat network goes down, it could quickly repair itself without intervention of central control nodes. Therefore the system becomes more predictable because signaling mechanisms are way less dependent on the backhaul infrastructure and other (anchor) entities.

Another important benefit of UFA-HIP is the potential to prevent suboptimal routing situations and realize advanced resource efficiency. In a common hierarchical architecture, all traffic passes through the centralized anchor nodes, which likely increases the routing path and results in suboptimal traffic routing compared to the flat use-cases.

However, in order to exploit all the above benefits and advantages, some challenges that flat architectures face must be concerned.

Chapter 6
Conclusions and Future Research

Supporting localized or micromobility management, location privacy enhancement, network mobility management and distributed mobility management are very important scenarios for emerging application areas and use-cases in the all-IP world of modern telecommunications. In my dissertation I have presented new schemes, protocols and algorithms to support these scenarios, improve the performance of legacy solutions in these use-cases and such increase the quality of mobile applications in general.

The proposed anycast based micromobility management method and the introduced anycast domain planning scheme would easily represent a convenient hop-by-hop routing based but still scalable technique for mobile operators to deploy transparent and built-in micromobility management. However, my proposal requires the standardization and widespread implementation of IPv6 anycast routing and group management protocols. Supporting this standardization work is a possible future research direction, especially if protocol parameters defining the speed of routing convergence comes into picture. Note, that in case of ABMF routing convergence defines handover latency, therefore increasing convergence time is an essential goal for ABMF supporting anyast routing solutions. Also important to solve security and secure group management issues of IPv6 anycasting: without such a technique, routing information injection into the routing system is a potential threat.

The introduced location privacy aware domain planning algorithms would provide an opportunity for mobile operators to configure micromobility domains and define gateway placement policies in distributed architectures in a way which guarantees a near optimum tradeoff between the registration signaling load and the paging cost, while also maximizing the domain's location privacy capabilities based on different considerations. Domain planning is an important issue in the design of future's highly distributed or even flat mobile networks, since IP address changes will occur much frequently in such architectures, therefore structure of domains will show even more serious impacts on IP-dependent location privacy and also on the total mobility management cost of mobile nodes. Of course operators will only apply such techniques if the user awareness for location privacy protection will increase above a certain level, such forcing the application of strong privacy enhancing technologies even during the network planning phase. In order to provide a more general and complete domain planning scheme, it is advisable to combine the features of the proposed algorithm variants and integrate them in a complex and more adaptive design solution. The proposed concept of location privacy aware network planning fits into the topic of personal paging area design and could help to create novel set of services for power users wanting to pay more money for advanced network services such network-aided enhanced IP location privacy. Optimization of the proposed techniques for heterogeneous integrated Wi-Fi Femto (IFW) architectures is also a promising future research direction.

My proposed, location information aided predictive mobility management framework and handover scheme extends the standardized NEMO BS solution for network mobility, and combines the benefits of MCoA with a new prediction-driven cross-layer management entity. I have shown that with an appropriate setup the prediction engine will not suffer from the errors of wrong positioning measurements, which makes my proposed system a solid, trustworthy and practical extension of NEMO BS in multihomed configurations. The scheme was successfully applied in the BOSS project [155] and served as the main mobility management solution for the on board wireless secured video surveillance system designed by the consortium for railways. A natural and practical enhancement of the proposal could be the integration of the solution with IEEE 802.21 MIH and/or ANDSF functions. Also an

interesting future research topic is to analyze the adaptation possibilities of the proposed technique into ITS/C-ITS system architectures under standardization. The C-ITS concept of Local Dynamic Map (LDM) could accommodate several modules of my proposal and could provide a standardized and easy-to-use toolset to implement and further improve my MCoA based, GNSS aided predictive handover management scheme.

The separation of locator and identifier information is probably one of the main evolution trends of the future Internet. The Host Identity Protocol is a security control protocol providing true, cryptographic ID/Loc separation, IP-mobility and multihoming. In HIP-enabled nodes, applications use persistent host identities instead of IP addresses for addressing. Any type of mobility is hidden from the application and transport layer. In current 3GPP networks, non-3GPP access is protected by IKEv2 and IPsec protocols. HIP could replace IKEv2 currently defined for non-3GPP access as network access security protocol, if it performs better in L3 re-authentication and IPsec security association establishment procedures. Seamless inter-system handover between non-3GPP accesses is not covered by current 3GPP standards, however HIP could also support seamless inter-system handover between non-3GPP access networks. In distributed or flat architectures, containing multiple distributed P-GWs, intra-3GPP mobility will lead to frequent inter-PGW handovers. That is the reason why my proposed HIP based micromobiliy and UFA-HIP distributed mobility could play an essential role in future mobile Internet architectures, and also this is the motivation to provide NEMO support also in the host identity layer by my introduced HIP-NEMO scheme. However, introduction of HIP technologies in current or evolving mobile architectures is not an easy job: the structural modifications inside the common TCP/IP protocol stack raises serious deployment concerns which should be tackled for widespread application of HIP based networking solutions. In case of HIP delegation-based services, support of non-HIP enabled peers can be solved by example using HIP proxies [156]. My HIP-based schemes were successfully applied in project MEVICO [157] as the building blocks of a possible green-field alternative to support mobile networks evolution towards a more distributed architecture and enhanced individual communications experience. An important future research direction in this topic is the analysis of cooperation opportunities between HIP based advanced mobility solutions and Software Defined Mobile Networks (SDMNs) where traffic driven dynamic reconfiguration and optimization of radio, transport and core network resources are managed using centralized and automated controlling capabilities and open interfaces. Secure SDMN signaling and support of complex mobility scenarios are just two possible application areas of HIP techniques in software defined networks. The proliferation of softwarized, virtualized and cloudified mobile Internet infrastructures also require reconciliation of mobility solutions in general as the foreseen real-time traffic optimization in SDMNs creates a new paradigm with endless possibilities for handover management.

References

[1] CISCO, “Global mobile data traffic forecast update, 2013–2018.” Tech. rep., Cisco VNI White Paper, Feb-2014.

[2] UMTS forum, “UMTS Forum Report 44, Mobile traffic forecasts 2010-2020 report: Recognising the promise of mobile broadband,” UMTS Forum White Paper, Jan. 2011.

[3] UMTS forum, “UMTS Forum, REPORT NO 37, Magic Mobile Future 2010-2020,” UMTS Forum White Paper, Apr. 2005.

[4] C. Antón-Haro, T. Lestable, Y. Lin, N. Nikaein, T. Watteyne, and J. Alonso-Zarate, “Machine-to-machine: an emerging communication paradigm,” *Transactions on Emerging Telecommunications Technologies*, vol. 24, no. 4, pp. 353–354, 2013.

[5] C. Perkins, *IP Mobility Support for IPv4*. IETF, 2002.

[6] C. Perkins, D. Johnson, and J. Arkko, *Mobility Support in IPv6*. IETF, 2011.

[7] D. E. 3rd, J. Reagle, and D. Solo, *(Extensible Markup Language) XML-Signature Syntax and Processing*. IETF, 2002.

[8] S. J. Koh, M. J. Chang, and M. Lee, “mSCTP for soft handover in transport layer,” *Communications Letters, IEEE*, vol. 8, no. 3, pp. 189–191, Mar. 2004.

[9] H. Soliman, C. Castelluccia, K. ElMalki, and L. Bellier, *Hierarchical Mobile IPv6 (HMIPv6) Mobility Management*. IETF, 2008.

[10] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, *Proxy Mobile IPv6*. IETF, 2008.

[11] A. G. Valkó, “Cellular IP: A New Approach to Internet Host Mobility,” *SIGCOMM Comput. Commun. Rev.*, vol. 29, no. 1, pp. 50–65, Jan. 1999.

[12] T. Ernst, “The Information Technology Era of the Vehicular Industry,” *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 2, pp. 49–52, Apr. 2006.

[13] A. Stevens and J. Hopkin, “Benefits and deployment opportunities for vehicle/roadside cooperative ITS,” in *Road Transport Information and Control (RTIC 2012), IET and ITS Conference on*, 2012, pp. 1–6.

[14] L. A. DaSilva, S. F. Midkiff, J. S. Park, G. C. Hadjichristofi, I. Davis, N.J., K. S. Phanse, and T. Lin, “Network mobility and protocol interoperability in ad hoc networks,” *Communications Magazine, IEEE*, vol. 42, no. 11, pp. 88–96, Nov. 2004.

[15] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert, *Network Mobility (NEMO) Basic Support Protocol*. IETF, 2005.

[16] M. Bauer, P. Bosch, N. Khrais, L. G. Samuel, and P. Schefczik, “The UMTS base station router,” *Bell Labs Technical Journal*, vol. 11, no. 4, pp. 93–111, 2007.

[17] K. Daoud, P. Herbelin, and N. Crespi, “UFA: Ultra Flat Architecture for high bitrate services in mobile networks,” in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, 2008, pp. 1–6.

[18] ITU-T, “General requirements for ID/locator separation in NGN, ITU-T Y.2015 (Y.ipsplit).” ITU-T Draft Recommendation, 06-Feb-2009.

[19] V. P. Kafle, H. Otsuki, and M. Inoue, “An ID/locator split architecture for future networks,” *Communications Magazine, IEEE*, vol. 48, no. 2, pp. 138–144, Feb. 2010.

[20] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson, *Host Identity Protocol*. IETF, 2008.

[21] R. Moskowitz and P. Nikander, *Host Identity Protocol (HIP) Architecture*. IETF, 2006.

[22] P. Nikander, T. Henderson, C. Vogt, and J. Arkko, *End-Host Mobility and Multihoming with the Host Identity Protocol*. IETF, 2008.

[23] J. Laganier and L. Eggert, *Host Identity Protocol (HIP) Rendezvous Extension*. IETF, 2008.

[24] X. He, D. Funato, and T. Kawahara, “A dynamic micromobility domain construction scheme,” in *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, 2003, vol. 3, pp. 2495–2499 vol.3.

[25] V. Simon and S. Imre, “A Simulated Annealing Based Location Area Optimization in Next Generation Mobile Networks,” *Mob. Inf. Syst.*, vol. 3, no. 3,4, pp. 221–232, Dec. 2007.

[26] S. Pack, M. Nam, and Y. Choi, “A study on optimal hierarchy in multi-level hierarchical mobile IPv6 networks,” in *Global Telecommunications Conference, 2004. GLOBECOM ’04. IEEE*, 2004, vol. 2, pp. 1290–1294 Vol.2.

[27] S.-W. Lo, T.-W. Kuo, K.-Y. Lam, and G.-H. Li, “Efficient location area planning for cellular networks with hierarchical location databases,” *Computer Networks*, vol. 45, no. 6, pp. 715 – 730, 2004.

[28] E. Cayirci and I. F. Akyildiz, “Optimal location area design to minimize registration signaling traffic in wireless systems,” *Mobile Computing, IEEE Transactions on*, vol. 2, no. 1, pp. 76–85, Jan. 2003.

[29] N. Montavont, T. Noel, and T. Ernst, “Multihoming in nested mobile networking,” in *Applications and the Internet Workshops, 2004. SAINT 2004 Workshops. 2004 International Symposium on*, 2004, pp. 184–189.

[30] K. Park, S. Han, and J. Song, “Selective Handover Technique on Multihomed Mobile Network Environment,” in *Computational Science – ICCS 2006*, vol. 3992, V. Alexandrov, G. Albada, P. A. Sloot, and J. Dongarra, Eds. Springer Berlin Heidelberg, 2006, pp. 1081–1088.
[31] V. P. Kafle, E. Kamioka, and S. Yamada, “MoRaRo: Mobile Router-Assisted Route Optimization for Network Mobility (NEMO) Support,” *IEICE - Trans. Inf. Syst.*, vol. E89-D, no. 1, pp. 158–170, Jan. 2006.
[32] M. Calderon, C. J. Bernardos, M. Bagnulo, I. Soto, and A. De La Oliva, “Design and Experimental Evaluation of a Route Optimization Solution for NEMO,” *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 9, pp. 1702–1716, Sep. 2006.
[33] M. Jeong, Y. Cho, and J. Park, “Hierarchical mobile network binding scheme for route optimization in NEMO,” *Wireless Personal Communications*, vol. 43, no. 3, pp. 975–995, 2007.
[34] T. K. Tan and A. Samsudin, “Efficient NEMO security management via CAP-KI,” in *Telecommunications and Malaysia International Conference on Communications, 2007. ICT-MICC 2007. IEEE International Conference on*, 2007, pp. 140–144.
[35] L.-S. Li, S.-S. Tzeng, R.-C. Bai, and M.-T. Li, “End to End Security and Path Security in Network Mobility,” in *Parallel Processing Workshops (ICPPW), 2011 40th International Conference on*, 2011, pp. 16–21.
[36] Y.-H. Wang, K.-F. Huang, and H.-Y. Ho, “A Seamless Handover Scheme with Pre-registration in NEMO,” in *Advanced Information Networking and Applications Workshops, 2009. WAINA ’09. International Conference on*, 2009, pp. 338–344.
[37] C.-W. Lee, Y. S. Sun, and M.-C. Chen, “HiMIP-NEMO: Combining Cross-Layer Network Mobility Management and Resource Allocation for Fast QoS-Handovers,” in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, 2008, pp. 2282–2286.
[38] H.-D. Park, Y.-H. Kwon, K.-W. Lee, Y.-S. Choi, S.-H. Lee, and Y.-Z. Cho, “Network Mobility Management Using Predictive Binding Update,” in *Distributed Computing – IWDC 2005*, vol. 3741, A. Pal, A. Kshemkalyani, R. Kumar, and A. Gupta, Eds. Springer Berlin Heidelberg, 2005, pp. 560–565.
[39] J. Montavont, J. Lorchat, and T. Noel, “Deploying NEMO: a Practical Approach,” in *ITS Telecommunications Proceedings, 2006 6th International Conference on*, 2006, pp. 1053–1056.
[40] K. Lan, E. Perera, H. Petander, C. Dwertmann, L. Libman, and M. Hassan, “MOBNET: the design and implementation of a network mobility testbed for NEMO protocol,” in *Local and Metropolitan Area Networks, 2005. LANMAN 2005. The 14th IEEE Workshop on*, 2005, p. 6 pp.–6.
[41] C.-M. Huang, C.-H. Lee, and J.-R. Zheng, “A Novel SIP-Based Route Optimization for Network Mobility,” *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 9, pp. 1682–1691, Sep. 2006.
[42] P. Bertin, S. Bonjour, and J.-M. Bonnin, “Distributed or Centralized Mobility?,” in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009, pp. 1 –6.
[43] K. Daoud, P. Herbelin, K. Guillouard, and N. Crespi, “Performance and implementation of UFA: A SIP-based Ultra Flat mobile network architecture,” in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, 2009, pp. 793–797.
[44] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, *SIP: Session Initiation Protocol*. IETF, 2002.
[45] V. Simon and S. Imre, “Location Area Design Algorithms for Reducing Signalling Overhead in Mobile Networks,” in *3rd Inter. Conf. On Advances in Mobile Multimedia, MoMM’05*, Kuala Lumpur, Malaysia, 2005, pp. 365–375.
[46] A. Varga and R. Hornig, “An Overview of the OMNeT++ Simulation Environment,” in *Proceedings of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems & Workshops*, ICST, Brussels, Belgium, Belgium, 2008, pp. 60:1–60:10.
[47] P. Reinbold and O. Bonaventure, “IP micro-mobility protocols,” *Communications Surveys Tutorials, IEEE*, vol. 5, no. 1, pp. 40–57, Third 2003.
[48] R. Ramjee, K. Varadhan, L. Salgarelli, S. R. Thuel, S.-Y. Wang, and T. La Porta, “HAWAII: a domain-based approach for supporting mobility in wide-area wireless networks,” *Networking, IEEE/ACM Transactions on*, vol. 10, no. 3, pp. 396–410, Jun. 2002.
[49] R. Ramjee, L. Li, T. La Porta, and S. Kasera, “IP Paging Service for Mobile Hosts,” *Wireless Networks*, vol. 8, no. 5, pp. 427–441, 2002.
[50] C. Partridge, T. Mendez, and W. Milliken, *Host Anycasting Service*. IETF, 1993.
[51] S. Doi, S. Ata, H. Kitamura, and M. Murata, “IPv6 anycast for simple and effective service-oriented communications,” *Communications Magazine, IEEE*, vol. 42, no. 5, pp. 163–171, May 2004.
[52] S. Matsunaga, S. Ata, H. Kitamura, and M. Murata, “Applications of IPv6 Anycasting.” IETF Internet Draft, Feb-2005.
[53] R. Hinden and S. Deering, *IP Version 6 Addressing Architecture*. IETF, 2006.

[54] M. Hashimoto, S. Ata, H. Kitamura, and M. Murata, “IPv6 Anycast Terminolgy Definition.” IETF Internet Draft, Jan-2006.
[55] S. Ata, H. Kitamura, and M. Murata, “Possible Deployment Scenarios for IPv6 Anycasting.” IETF Internet Draft, Oct-2004.
[56] S. Matsunaga, S. Ata, H. Kitamura, and M. Murata, “Design and implementation of IPv6 anycast routing protocol: PIA-SM,” in *Advanced Information Networking and Applications, 2005. AINA 2005. 19th International Conference on*, 2005, vol. 2, pp. 839–844 vol.2.
[57] S. Doi, S. Ata, H. Kitamura, and M. Murata, “Design, implementation and evaluation of routing protocols for IPv6 anycast communication,” in *Advanced Information Networking and Applications, 2005. AINA 2005. 19th International Conference on*, 2005, vol. 2, pp. 833–838 vol.2.
[58] V. D. Park and J. P. Macker, “Anycast Routing for Mobile Services,” in *Conf. Info. Sci. and Sys*, Baltimore, 1999.
[59] V. D. Park and J. P. Macker, “Anycast routing for mobile networking,” in *Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE*, 1999, vol. 1, pp. 1–5 vol.1.
[60] D. Xuan, W. Jia, W. Zhao, and H. Zhu, “A routing protocol for anycast messages,” *Parallel and Distributed Systems, IEEE Transactions on*, vol. 11, no. 6, pp. 571–588, Jun. 2000.
[61] E. Shim, *Mobility Management in the Wireless Internet*, PhD Thesis. Columbia University, 2004.
[62] Z. Zhou, G. Xu, J. He, J. Jiang, and C. Deng, “Research of Secure Anycast Group Management,” in *Networked Computing and Advanced Information Management, 2008. NCM ’08. Fourth International Conference on*, 2008, vol. 1, pp. 604–608.
[63] R. Hinden and S. Deering, *Internet Protocol Version 6 (IPv6) Addressing Architecture*. IETF, 2003.
[64] V. Ponnusamy, E. K. Karuppiah, and R. Abdullah, “Anycast group membership management protocol,” in *Communications, 2003. APCC 2003. The 9th Asia-Pacific Conference on*, 2003, vol. 3, pp. 1052–1056 Vol.3.
[65] Y. Wang, L. Zhang, Z. Han, and W. Yan, “Anycast extensions to OSPFv3,” in *Parallel and Distributed Systems, 2005. Proceedings. 11th International Conference on*, 2005, vol. 1, pp. 223–229 Vol. 1.
[66] R. Coltun, D. Ferguson, J. Moy, and A. Lindem, *OSPF for IPv6*. IETF, 2008.
[67] L. Osborne, A. Abdel-Hamid, and R. Ramadugu, “A performance comparison of Mobile IPv6, hierarchical Mobile IPv6, and Mobile IPv6 regional registrations,” in *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, 2005, vol. 2, pp. 1545–1550 vol.2.
[68] J. G. Markoulidakis, G. L. Lyberopoulos, D. F. Tsirkas, and E. D. Sykas, “Evaluation of location area planning scenarios in future mobile telecommunication systems,” *Wireless Networks*, vol. 1, no. 1, pp. 17–29, 1995.
[69] S. Tabbane, “Location management methods for third generation mobile systems,” *Communications Magazine, IEEE*, vol. 35, no. 8, pp. 72–78, 83–4, Aug. 1997.
[70] V. Simon and S. Imre, “Location Area Design Algorithms for Minimizing Signalling Costs in Mobile Networks,” in *Mobile Computing: Concepts, Methodologies, Tools, and Applications*, IGI Global, 2009, pp. 682–695.
[71] N. B. Prajapati, R. R. Agravat, and M. I. Hasan, “Simulated Annealing for Location Area Planning in Cellular networks,” *International journal on applications of graph theory in wireless ad hoc networks and sensor networks (GRAPH-HOC)*, vol. 2, no. 1, pp. 1–7, Mar. 2010.
[72] G. Qian, L. Guang-xia, L. Jing, X. Yi-qun, and Z. Ming, “Location Area Design for GEO Mobile Satellite System,” in *Computer Engineering and Applications (ICCEA), 2010 Second International Conference on*, 2010, vol. 1, pp. 525–529.
[73] A. Kumar, M. N. Umesh, and R. Jha, “Mobility Modeling of Rush Hour Traffic for Location Area Design in Cellular Networks,” in *Proceedings of the 3rd ACM International Workshop on Wireless Mobile Multimedia*, New York, NY, USA, 2000, pp. 48–54.
[74] V. Simon and S. Imre, “A Domain Forming Algorithm for Next Generation, IP Based Mobile Networks,” in *SoftCOM2004*, Dubrovnik, Croatia, 2004, pp. 289–292.
[75] P. S. Bhattacharjee, D. Saha, and A. Mukherjee, “Heuristics for assignment of cells to switches in a PCSN: a comparative study,” in *Personal Wireless Communication, 1999 IEEE International Conference on*, 1999, pp. 331–334.
[76] P. J. M. Laarhoven and E. H. L. Aarts, Eds., *Simulated Annealing: Theory and Applications*. Norwell, MA, USA: Kluwer Academic Publishers, 1987.
[77] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by Simulated Annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
[78] J. Laganier, T. Koponen, and L. Eggert, *Host Identity Protocol (HIP) Registration Extension*. IETF, 2008.

[79] J. Ylitalo, J. Melén, P. Nikander, and V. Torvinen, “Re-thinking Security in IP Based Micro-Mobility,” in *Information Security*, vol. 3225, K. Zhang and Y. Zheng, Eds. Springer Berlin Heidelberg, 2004, pp. 318–329.
[80] ITU-T, “Framework of ID/LOC separation in IPv6-based NGN (Y.ipv6split).” ITU-T Draft Recommendation, 05-Mar-2009.
[81] P. Nikander and J. Arkko, “Delegation of Signalling Rights,” in *Security Protocols*, vol. 2845, B. Christianson, B. Crispo, J. Malcolm, and M. Roe, Eds. Springer Berlin Heidelberg, 2004, pp. 203–214.
[82] S. Herborn, A. Huber, R. Boreli, and A. Seneviratne, “Secure Host Identity Delegation for Mobility,” in *Communication Systems Software and Middleware, 2007. COMSWARE 2007. 2nd International Conference on*, 2007, pp. 1–9.
[83] P. Jokela, R. Moskowitz, and P. Nikander, *Using the Encapsulating Security Payload (ESP) Transport Format with the Host Identity Protocol (HIP)*. IETF, 2008.
[84] P. Nikander and J. Laganier, *Host Identity Protocol (HIP) Domain Name System (DNS) Extensions*. IETF, 2008.
[85] P. Jokela, J. Melen, and J. Ylitalo, “HIP Service Discovery.” IETF Internet Draft, Jun-2006.
[86] Z. Kovacshazi and R. Vida, “Host Identity Specific Multicast,” in *Networking and Services, 2007. ICNS. Third International Conference on*, 2007, pp. 1–1.
[87] “The INET Framework for OMNeT++,” Mar-2014. [Online]. Available: http://inet.omnetpp.org/.
[88] S. Kent and R. Atkinson, *IP Encapsulating Security Payload (ESP)*. IETF, 1998.
[89] “Infrastructure for HIP (InfraHIP): Project focusing on developing the missing infrastructure pieces of HIP,” Mar-2014. [Online]. Available: http://infrahip.hiit.fi/.
[90] A. R. Beresford and F. Stajano, “Location privacy in pervasive computing,” *Pervasive Computing, IEEE*, vol. 2, no. 1, pp. 46–55, Jan. 2003.
[91] A. Lakhina, J. W. Byers, M. Crovella, and I. Matta, “On the geographic location of Internet resources,” *Selected Areas in Communications, IEEE Journal on*, vol. 21, no. 6, pp. 934–948, Aug. 2003.
[92] M. J. Freedman, M. Vutukuru, N. Feamster, and H. Balakrishnan, “Geographic Locality of IP Prefixes,” in *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement*, Berkeley, CA, USA, 2005, pp. 13–13.
[93] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, “Constraint-Based Geolocation of Internet Hosts,” *Networking, IEEE/ACM Transactions on*, vol. 14, no. 6, pp. 1219–1232, Dec. 2006.
[94] B. Eriksson, P. Barford, J. Sommers, and R. Nowak, “A Learning-Based Approach for IP Geolocation,” in *Passive and Active Measurement*, vol. 6032, A. Krishnamurthy and B. Plattner, Eds. Springer Berlin Heidelberg, 2010, pp. 171–180.
[95] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang, “Towards Street-level Client-independent IP Geolocation,” in *Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation*, Berkeley, CA, USA, 2011, pp. 27–27.
[96] R. Koodli, *IP Address Location Privacy and Mobile IPv6: Problem Statement*. IETF, 2007.
[97] W. Haddad, E. Nordmark, F. Dupont, and M. Bagnulo, “Anonymous Identifiers (ALIEN): Privacy Threat Model for Mobile and Multi-Homed Nodes.” IETF Internet Draft, Jun-2006.
[98] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, “Femtocell networks: a survey,” *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59–67, Sep. 2008.
[99] C. Díaz, S. Seys, J. Claessens, and B. Preneel, “Towards Measuring Anonymity,” in *Proceedings of the 2Nd International Conference on Privacy Enhancing Technologies*, Berlin, Heidelberg, 2003, pp. 54–68.
[100] A. Serjantov and G. Danezis, “Towards an Information Theoretic Metric for Anonymity,” in *Proceedings of the 2Nd International Conference on Privacy Enhancing Technologies*, Berlin, Heidelberg, 2003, pp. 41–53.
[101] R. Shokri, J. Freudiger, M. Jadliwala, and J.-P. Hubaux, “A Distortion-based Metric for Location Privacy,” in *Proceedings of the 8th ACM Workshop on Privacy in the Electronic Society*, New York, NY, USA, 2009, pp. 21–30.
[102] M. Li, K. Sampigethaya, L. Huang, and R. Poovendran, “Swing & Swap: User-centric Approaches Towards Maximizing Location Privacy,” in *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society*, New York, NY, USA, 2006, pp. 19–28.
[103] T. Jiang, H. J. Wang, and Y.-C. Hu, “Preserving Location Privacy in Wireless Lans,” in *Proceedings of the 5th International Conference on Mobile Systems, Applications and Services*, New York, NY, USA, 2007, pp. 246–257.
[104] L. Huang, H. Yamane, K. Matsuura, and K. Sezaki, “Silent Cascade: Enhancing Location Privacy Without Communication QoS Degradation,” in *Proceedings of the Third International Conference on Security in Pervasive Computing*, Berlin, Heidelberg, 2006, pp. 165–180.

[105] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, “Preserving Privacy in Gps Traces via Uncertainty-aware Path Cloaking,” in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, New York, NY, USA, 2007, pp. 161–171.
[106] B. Hoh, M. Gruteser, R. Herring, J. Ban, D. Work, J.-C. Herrera, A. M. Bayen, M. Annavaram, and Q. Jacobson, “Virtual Trip Lines for Distributed Privacy-preserving Traffic Monitoring,” in *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*, New York, NY, USA, 2008, pp. 15–28.
[107] J. T. Meyerowitz and R. R. Choudhury, “Realtime Location Privacy via Mobility Prediction: Creating Confusion at Crossroads,” in *Proceedings of the 10th Workshop on Mobile Computing Systems and Applications*, New York, NY, USA, 2009, pp. 2:1–2:6.
[108] R. Wakikawa, V. Devarapalli, G. Tsirtsis, T. Ernst, and K. Nagami, *Multiple Care-of Addresses Registration*. IETF, 2009.
[109] G. Tsirtsis, H. Soliman, N. Montavont, G. Giaretta, and K. Kuladinithi, *Flow Bindings in Mobile IPv6 and Network Mobility (NEMO) Basic Support*. IETF, 2011.
[110] G. Tsirtsis, G. Giarreta, H. Soliman, and N. Montavont, *Traffic Selectors for Flow Bindings*. IETF, 2011.
[111] R. Koodli, *Fast Handovers for Mobile IPv6*. IETF, 2005.
[112] S. S. Wang and C.-H. Wu, “Effective handoff method using mobile location information,” in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, 2001, vol. 4, pp. 2585–2589 vol.4.
[113] A. Dutta, S. Chakravarty, K. Taniuchi, V. Fajardo, Y. Ohba, D. Famolari, and H. Schulzrinne, “An Experimental Study of Location Assisted Proactive Handover,” in *Global Telecommunications Conference, 2007. GLOBECOM ’07. IEEE*, 2007, pp. 2037–2042.
[114] S. Thomson, T. Narten, and T. Jinmei, *IPv6 Stateless Address Autoconfiguration*. IETF, 2007.
[115] R. Droms, J. Bound, B. Volz, T. Lemon, C. Perkins, and M. Carney, *Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*. IETF, 2003.
[116] M. Han, K.-S. Han, and D.-J. Lee, “Fast IP Handover Performance Improvements Using Performance Enhancing Proxys between Satellite Networks and Wireless LAN Networks for High-Speed Trains,” in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, 2008, pp. 2341–2344.
[117] C. Ng, P. Thubert, M. Watari, and F. Zhao, *Network Mobility Route Optimization Problem Statement*. IETF, 2007.
[118] C. Ng, F. Zhao, M. Watari, and P. Thubert, *Network Mobility Route Optimization Solution Space Analysis*. IETF, 2007.
[119] S. Jung, F. Zhao, S. F. Wu, and H. Kim, “Threat Analysis on NEtwork MObility (NEMO),” in *Information and Communications Security*, vol. 3269, J. Lopez, S. Qing, and E. Okamoto, Eds. Springer Berlin Heidelberg, 2004, pp. 331–342.
[120] A. Deleplace, T. Ernst, and T. Noel, “Multihoming in Nested Mobile Networks with Route Optimization,” in *Applications and the Internet Workshops, 2007. SAINT Workshops 2007. International Symposium on*, 2007, pp. 49–49.
[121] M. Kim and K. Chae, “A Fast Defense Mechanism Against IP Spoofing Traffic in a NEMO Environment,” in *Proceedings of the 2005 International Conference on Information Networking: Convergence in Broadband and Mobile Networking*, Berlin, Heidelberg, 2005, pp. 843–852.
[122] T. Oiwa, M. Kunishi, M. Ishiyama, M. Kohno, and F. Teraoka, “A network mobility protocol based on LIN6,” in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, 2003, vol. 3, pp. 1984–1988 Vol.3.
[123] J. Ylitalo, “Re-thinking Security in Network Mobility,” in *NDSS Wireless and Security Workshop*, San Diego, CA, USA, 2005.
[124] S. Herborn, L. Haslett, R. Boreli, and A. Seneviratne, “HarMoNy - HIP Mobile Networks,” in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, 2006, vol. 2, pp. 871–875.
[125] R. Hinden, S. Deering, and E. Nordmark, *IPv6 Global Unicast Address Format*. IETF, 2003.
[126] H. Holma, A. Toskala, K. Ranta-aho, and J. Pirskanen, “High-Speed Packet Access Evolution in 3GPP Release 7 [Topics in Radio Communications],” *Communications Magazine, IEEE*, vol. 45, no. 12, pp. 29–35, Dec. 2007.
[127] M. Corici, D. Vingarzan, T. Magedanz, and T. Magedanz, “3GPP Evolved Packet Core - the Mass Wireless Broadband all-IP architecture,” in *Telecommunications: The Infrastructure for the 21st Century (WTC), 2010*, 2010, pp. 1–6.
[128] A. Khandekar, N. Bhushan, J. Tingfang, and V. Vanghi, “LTE-Advanced: Heterogeneous networks,” in *Wireless Conference (EW), 2010 European*, 2010, pp. 978–982.
[129] D. Niyato and E. Hossain, “WIRELESS BROADBAND ACCESS: WIMAX AND BEYOND - Integration of WiMAX and WiFi: Optimal Pricing for Bandwidth Sharing,” *Communications Magazine, IEEE*, vol. 45, no. 5, pp. 140–146, May 2007.

[130] P. Agrawal, J.-H. Yeh, J.-C. Chen, and T. Zhang, "IP multimedia subsystems in 3GPP and 3GPP2: overview and scalability issues," *Communications Magazine, IEEE*, vol. 46, no. 1, pp. 138–145, Jan. 2008.
[131] CISCO, "Cisco Visual Networking Index: The Zettabyte Era-Trends and Analysis." Cisco White Paper, May-2013.
[132] Berg Insight, "M2M Research Series: The Global Wireless M2M Market." Report (Fourth Edition), 2011.
[133] K. Daoud, P. Herbelin, and N. Crespi, "One-node-based mobile architecture for a better QoS control," in *Wireless Days, 2008. WD '08. 1st IFIP*, 2008, pp. 1–5.
[134] IEEE, "IEEE Standard for Local and metropolitan area networks- Part 21: Media Independent Handover." IEEE Std 802.21-2008, 2009.
[135] J. Loughney, M. Nakhjiri, C. Perkins, and R. Koodli, *Context Transfer Protocol (CXTP)*. IETF, 2005.
[136] J. Ylitalo, P. Salmela, and H. Tschofenig, "SPINAT: Integrating IPsec into Overlay Routing," in *Security and Privacy for Emerging Areas in Communications Networks, 2005. SecureComm 2005. First International Conference on*, 2005, pp. 315–326.
[137] T. Melia, G. Bajko, S. Das, N. Golmie, and J. Zuniga, *IEEE 802.21 Mobility Services Framework Design (MSFD)*. IETF, 2009.
[138] A. Gurtov, D. Korzun, A. Lukyanenko, and P. Nikander, "Hi3: An Efficient and Secure Networking Architecture for Mobile Hosts," *Comput. Commun.*, vol. 31, no. 10, pp. 2457–2467, Jun. 2008.
[139] H. Soliman, *Mobile IPv6 Support for Dual Stack Hosts and Routers*. IETF, 2009.
[140] A. Grilo, P. Estrela, and M. Nunes, "Terminal independent mobility for IP (TIMIP)," *Communications Magazine, IEEE*, vol. 39, no. 12, pp. 34–41, Dec. 2001.
[141] T. Melia, A. de la Oliva, A. Vidal, I. Soto, D. Corujo, and R. Aguiar, "Toward IP Converged Heterogeneous Mobility: A Network Controlled Approach," *Comput. Netw.*, vol. 51, no. 17, pp. 4849–4866, Dec. 2007.
[142] 3GPP, "Proxy Mobile IPv6 (PMIPv6) based Mobility and Tunneling protocols." TS 29.275, Dec-2010.
[143] 3GPP, "Mobility management based on Dual-Stack Mobile IPv6." TS 24.303, Dec-2010.
[144] 3GPP, "Architecture enhancements for non-3GPP accesses." TS 23.402, Jan-2011.
[145] 3GPP, "Local IP Access and Selected IP Traffic Offload." TS 23.829, May-2007.
[146] R.Wakikawa, R. Kuntz, Z. Thu, and L. Zhang, *Global HA to HA Protocol Specification*. IETF Draft draft-wakikawa-mext-global-haha-spec-02, 2011.
[147] M. Fischer, F.-U. Andersen, A. Kopsel, G. Schafer, and M. Schlager, "A Distributed IP Mobility Approach for 3G SAE," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, 2008, pp. 1–6.
[148] M. Fischer, F.-U. Andersen, A. Kopsel, G. Schafer, and M. Schlager, "A Distributed IP Mobility Approach for 3G SAE," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, 2008, pp. 1 –6.
[149] R. Farha, K. Khavari, N. Abji, and A. Leon-Garcia, "Peer-to-Peer Mobility Management for all-IP Networks," in *Communications, 2006. ICC '06. IEEE International Conference on*, 2006, vol. 5, pp. 1946–1952.
[150] L. Yu, Z. Zhijun, L. Tao, and T. Hui, "Distributed mobility management based on flat network architecture," in *Wireless Internet Conference (WICON), 2010 The 5th Annual ICST*, 2010, pp. 1–6.
[151] J. Arkko, C. Vogt, and W. Haddad, *Enhanced Route Optimization for Mobile IPv6*. IETF, 2007.
[152] A. C. Snoeren and H. Balakrishnan, "An End-to-end Approach to Host Mobility," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2000, pp. 155–166.
[153] R. Stewart, *Stream Control Transmission Protocol*. IETF, 2007.
[154] Z. Cao, B. He, Y. Shi, Q. Wu, and G. Zorn, *EAP Extensions for the EAP Re-authentication Protocol (ERP)*. IETF, 2012.
[155] "EUREKA Celtic-BOSS : On Board Wireless Secured Video Surveillance," Mar-2014. [Online]. Available: http://celtic-boss.mik.bme.hu.
[156] J. Melen, J. Ylitalo, and H. Salmela, "Host Identity Protocol-based Mobile Proxy." IETF Internet Draft, Aug-2009.
[157] "EUREKA-Celtic Plus MEVICO project: Mobile Networks Evolution for Individual Communications Experience," Mar-2014. [Online]. Available: http://www.mevico.org/.

Publications

Journal papers

[J1] László Bokor, László Lois, Csaba A. Szabó, Sándor Szabó: „A Novel Vertical Handover Mechanism for Media Streaming in Heterogeneous Wireless Architectures", Híradástechnika, English Issue of Selected Papers, Volume LXII. 2007/7. pp. 52-59. 2007.

[J2] Szabolcs Nováczki, László Bokor, Gábor Jeney, Sándor Imre: „Design and Evaluation of a Novel HIP-Based Network Mobility Protocol", Journal of Networks, Academy Publisher, ISSN: 1796-2056, Volume 3, Issue 1, pp. 10-24, January 2008.

[J3] Vilmos Simon, László Bokor, Sándor Imre: „A Hierarchical Network Design Solution for Mobile IPv6", Journal of Mobile Multimedia (JMM) © Rinton Press, ISSN: 1550-4646, Vol. 5, No.4 (December 2009) pp. 317-332, 2009.

[J4] László Bokor, Árpád Huszák, Gábor Jeney: „Novel Results on SCTP Multihoming Performance in Native IPv6 UMTS–WLAN Environments", International Journal of Communication Networks and Distributed Systems (IJCNDS), 2010 - Vol. 5, No.1/2 pp. 25 – 45. ISSN (Online/Print): 1754-3924/1754-3916, DOI: 10.1504/IJCNDS.2010.033966, 2010.

[J5] László Bokor, Zoltán Kanizsai, Gábor Jeney: „IMS-centric Evaluation of IPv4/IPv6 Transition Methods in 3G UMTS Systems", International Journal on Advances in Networks and Services, © IARIA, ISSN: 1942-2644, vol. 3, no. 3 & 4, pp. 402 – 416, 2010.

[J6] Zoltán Faigl, László Bokor, Pedro Miguel Neves, Khadija Daoud, Philippe Herbelin: „Evaluation of two integrated signalling schemes for the Ultra Flat Architecture using SIP, IEEE 802.21, and HIP/PMIP protocols", Computer Networks, Elsevier B.V., ISSN: 1389-1286, DOI: doi:10.1016/j.comnet.2011.02.005, 2011.

[J7] László Bokor, Zoltán Faigl, Jochen Eisl, Gerd Windisch, Components for Integrated Traffic Management – The MEVICO Approach, INFOCOMMUNICATIONS JOURNAL 3:(4) pp. 38-49. 2011.

[J8] László Bokor, Vilmos Simon, Sándor Imre, Evaluation of the Location Privacy Aware Micromobility Domain Planning Scheme, INFOCOMMUNICATIONS JOURNAL III:(3) pp. 38-49. 2011.

[J9] László Bokor, Zoltán Faigl, Sándor Imre, Flat Architectures: Towards Scalable Future Internet Mobility, LECTURE NOTES IN COMPUTER SCIENCE 6656: pp. 35-50. 2011.

[J10] Zoltán Kanizsai, László Bokor, Gábor Jeney: „An Anycast based Feedback Aggregation Scheme for Efficient Network Transparency in Cross-layer Design", PERIODICA POLYTECHNICA-ELECTRICAL ENGINEERING 55:(1-2) pp. 45-52. 2011.

[J11] András Takács, László Bokor: „A Distributed Dynamic Mobility Architecture with Integral Cross-Layered and Context-Aware Interface for Reliable Provision of High Bitrate mHealth Services", LECTURE NOTES OF THE INSTITUTE FOR COMPUTER SCIENCES SOCIAL-INFORMATICS AND TELECOMMUNICATIONS ENGINEERING 61: pp. 369-379. 2013.

[J12] Zoltán Faigl, László Bokor, Jani Pellikka, Andrei Gurtov: „Suitability analysis of existing and new authentication methods for future 3GPP Evolved Packet Core", COMPUTER NETWORKS 57:(17) pp. 3370-3388. 2013.

[J13] Zoltán Faigl, Jani Pellikka, László Bokor, Andrei Gurtov: „Performance evaluation of current and emerging authentication schemes for future 3GPP network architectures", COMPUTER NETWORKS 60: pp. 60-74. Paper COMPNW_5170. 2014.

[J14] László Bokor, Zoltán Faigl, Sándor Imre: „Survey and Evaluation of Advanced Mobility Management Schemes in the Host Identity Layer", International Journal of Wireless Networks and Broadband Technologies (IJWNBT), ISSN 2155-6261, 3(1), 34-59, January-March 2014.

[J15] László Bokor, József Kovács, Szabó Csaba Attila: „A Home Agent Initiated Handover Solution for Fine-grained Offloading in Future Mobile Internet Architectures: A Survey and Experimental Evaluation", under press, 2014.

[J16] László Bokor, Gábor Jeney, József Kovács: „A study on the performance of an advanced framework for prediction-based NEMO handovers in multihomed scenarios", submitted, 2014.

[J17] László Bokor, Zoltán Faigl, Sándor Imre: „An extensive analysis of the Host Identity Protocol based Ultra Flat Architecture",submitted, 2014.

Hungarian journal papers

[J18] Bokor László, Szabó Sándor: „Multimédia szolgáltatás a következő generációs (NGN) hálózatokban", Magyar Távközlés (XVI), Budapest, 2005/4. szám. pp. 14-19., 2005.

[J19] Bokor László, Szabó Sándor: „Az IMS megjelenése és alkalmazása cellás mobil hálózatokban”, Híradástechnika, LXI. Évfolyam, pp. 11-19., 2006/10. szám.

[J20] Nováczki Szabolcs, Bokor László, Imre Sándor: „A Host Identity Protocol, avagy egy új internetarchitektúra alapjai”, Magyar Távközlés (XVII), ISSN: 0865-9648, pp. 20-25. Budapest, 2006/4.

[J21] Kara Péter András, Bokor László, Imre Sándor, A mérőalanyok prekoncepciói által okozott torzítások hatása 3G videotelefonálás QoE kiértékelési eredményeire, Híradástechnika LXVI.:(2011/4) pp. 22-28. 2011.

Book chapters

[B1] István Dudás, László Bokor, Sándor Imre: „Survey and Extension of Applications and Services in IPv6 Anycasting”, Encyclopedia of Mobile Computing & Commerce, Idea Group Inc., ISBN: 978-1-59904-002-8 (hardcover), 978-1-59904-003-5 (ebook), New York, NY, USA. 2007.

[B2] László Bokor, Zoltán Németh, István Dudás, Sándor Imre: „Novel Results on MBMS Service Provisioning in UTMS/WLAN Heterogeneous Architectures”, Handbook of Research in Mobile Multimedia, 2nd edition, Section IV: Mobile Networks, Chapter XXVIII, IGI Global, ISBN: 978-1-60566-046-2 (hardcover), Ismail Khalil Ibrahim (ed.), USA. Released on September 25. 2008.

[B3] Szabolcs Nováczki, László Bokor, Gábor Jeney, Sándor Imre: „Emerging Mobility Applications of Host Identity Protocol”, Next Generation Mobile Networks and Ubiquitous Computing, IGI Global, ISBN: 9781605662503 (ISBN13), 160566250X (ISBN10), 9781605662510 (EISBN13), DOI: 10.4018/978-1-60566-250-3.ch019, Samuel Pierre (ed.), USA. 2010.

[B4] László Bokor, Vilmos Simon and Sándor Imre: „A Location Privacy Aware Network Planning Algorithm for Micromobility Protocols”, in Simulated Annealing, Theory with Applications, Book edited by: Rui Chibante, ISBN: 978-953-307-134-3, pp.: 75-98, Publisher: Sciyo 2010.

[B5] László Bokor, Szabolcs Nováczki, Sándor Imre, Host Identity Protocol: „The Enabler of Advanced Mobility Management Schemes”, In: Katalin Tarnay, Gusztáv Adamis, Tibor Dulai (ed.) Advanced Communication Protocol Technologies: Solutions, Methods, and Applications. Hershey; New York: IGI Global, Information Science Reference, (ISBN: ISBN 978-1-60960-732-6) pp. 247-272., 2011.

[B6] László Bokor, Jeney Gábor: „IPv4 / IPv6 Coexistence and Transition: Concepts, Mechanisms and Trends”, In: Katalin Tarnay, Gusztáv Adamis, Tibor Dulai (ed.), Advanced Communication Protocol Technologies: Solutions, Methods, and Applications., Hershey ; New York: IGI Global, Information Science Reference, (ISBN: ISBN 978-1-60960-732-6), pp. 156-177. 2011.

[B7] József Kovács, László Bokor, Zoltán Kanizsai, Sándor Imre: „Review of Advanced Mobility Solutions for Multimedia Networking in IPv6”, In: Dimitris Kanellopoulos (ed.) Intelligent Multimedia Technologies for Networking Applications: Techniques and Tools., Hershey: IGI Global, Information Science Reference, pp. 25-47. ISBN: 9781466628335, 2013.

Conference papers

[C1] István Dudás, László Bokor, Gábor Bilek, Gábor Jeney, dr. Sándor Imre: „Examining anycast address supported mobility management using Mobile IPv6 testbed”, MELECON 2004 (2004.05.11-15.), INSPEC: 8180353, ISBN: 0-7803-8271-4, pp.555 – 558, Vol.2. Dubrovnik, Croatia. 2004.

[C2] László Bokor, István Dudás, dr. Sándor Imre, „Anycast-based Micromobility”, SoftCOM 2005, (2005.09.15-17), Split, Marina-Frapa, Croatia, pp.125-130.

[C3] László Bokor, István Dudás, Sándor Szabó, dr. Sándor Imre: „Anycast-based Micromobility: A New Solution for Micromobility Management in IPv6”, MoMM 2005 (2005.09.17-21), ISBN: 3-85403-195-5, pp.68-75, Malaysia, Kuala Lumpur, 2005.

[C4] Szabolcs Nováczki, László Bokor, dr. Sándor Imre: „Micromobility Support in HIP: Survey and Extension of Host Identity Protocol”, DOI: 10.1109/MELECON.2006.1653184, ISBN: 1-4244-0087-2, MELECON 2006 (2006.05.16-19.), Málaga, Spain, pp.651-654, Vol.1.

[C5] László Bokor, Nicolas Montavont, Paolo Di Francesco, Thierry Ernst, Tobias Hof, Jari Korva: „ANEMONE: A Pan-European Testbed to Validate IPv6 Mobility Technologies”, SAINT-WONEMO 2007 (2007. 01.15-19.), DOI: 10.1109/SAINT-W.2007.25, ISBN: 0-7695-2757-4, Hiroshima, Japan, pp.44-48.

[C6] Szabolcs Nováczki, László Bokor, Sándor Imre: „A HIP based Network Mobility Protocol”, SAINT-WONEMO 2007 (2007. 01.15-19.), DOI: 10.1109/SAINT-W.2007.8, ISBN: 0-7695-2757-4, INSPEC: 9352994, Hiroshima, Japan, pp.48-52.

[C7] Csaba A. Szabó, Sándor Szabó, László Bokor: „Design considerations of a novel media streaming architecture for heterogeneous access environment", BWAN 2006, (2006. 09.20) Alghero, Sardinia, Italy, ACM ICP Series; ISBN:1-59593-532-0, Vol. 196, Article No. 3.

[C8] László Bokor, László Lois, Csaba A. Szabó, Sándor Szabó: „Testbed of a Novel Media Streaming Architecture for Heterogeneous Wireless Environment", Tridentcom2007, (2007.05.21-23), ISBN: 1-4244-0739-7 , pp. 220-230. Orlando, Florida, USA.

[C9] László Bokor, Vilmos Simon, István Dudás, Sándor Imre: „Anycast Subnet Optimization for Efficient IPv6 Mobility Management", IEEE GIIS 2007, DOI: 10.1109/GIIS.2007.4404188, ISBN 978-1-4244-1376-8, pp. 187-190, Marrakesh, Morocco, 2-6. July, 2007.

[C10] Thierry Ernst, László Bokor, Antoine Boutet, Yoann Lopez: „An Open Network for Testing, Verification and Validation of IPv6-based ITS Components", ITST 2007, (2007.07.6-8), DOI: 10.1109/ITST.2007.4295901, ISBN: 1-4244-1178-5, pp. 1-6. Sophia Antipolis, France 2007.

[C11] László Bokor, Szabolcs Nováczki, Sándor Imre: „A Complete HIP based Framework for Secure Micromobility", 5th @WAS International Conference on Advances in Mobile Computing and Multimedia, MoMM2007, ISBN 978-3-85403-230-4, pp. 111-122., Jakarta, Indonesia, 3-5 December 2007.

[C12] Vilmos Simon, László Bokor, Sándor Imre: „Novel Network Design Algorithm for Optimizing Hierarchical Mobile IPv6", 5th @WAS International Conference on Advances in Mobile Computing and Multimedia, MoMM2007, ISBN 978-3-85403-230-4, pp. 123-133., Jakarta, Indonesia, 3-5 December 2007.

[C13] László Bokor, Zoltán Németh, István Dudás, Sándor Imre: „MBMS Service Provisioning in UTMS/WLAN Heterogeneous Architectures", 5th @WAS International Conference on Advances in Mobile Computing and Multimedia, MoMM2007, ISBN 978-3-85403-230-4, pp. 41-52., Jakarta, Indonesia, 3-5 December 2007.

[C14] László Bokor, Zoltán Kanizsai, dr. Sándor Imre: „Simple QoS Provisioning Framework for MBMS in all-IP UMTS Networks", IEEE CN: CFP08MEL-CDR, ISBN: 978-1-4244-1633-2, MELECON 2008 (2008.05.5-7.), Ajaccio, France, pp. 286-292.

[C15] Nicolas Montavont, Antoine Boutet, Tanguy Ropitault, Manabu Tsukada, Thierry Ernst, Jari Korva, Cesar Viho, László Bokor: „Anemone: A ready-to-go testbed for IPv6 compliant Intelligent Transport Systems", 8th International Conference on Intelligent Transport System Telecommunicaitons (ITST 2008), Print ISBN: 978-1-4244-2857-1, DOI: 10.1109/ITST.2008.4740262, pp. 228-233, Phuket, Thailand, October 22-24, 2008.

[C16] László Bokor, Árpád Huszák, Gábor Jeney: „On SCTP Multihoming Performance in Native IPv6 UMTS-WLAN Environments", In the proceedings of the Fifth International ICST Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom2009), ISBN: 978-1-4244-2847-2, Print ISBN: 978-1-4244-2846-5 DOI: 10.1109/TRIDENTCOM.2009.4976216 , Washington D.C., USA, April 06-08, 2009.

[C17] László Bokor, Szabolcs Nováczki, László Tamás Zeke, Gábor Jeney: „Design and Evaluation of Host Identity Protocol (HIP) Simulation Framework for INET/OMNeT++", in the proceedings of the 12-th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM 2009), ISBN:978-1-60558-616-8, DOI: 10.1145/1641804.1641827, pp. 124-133, Tenerife, Canary Islands, Spain, 26-30 October 2009.

[C18] László Bokor, László Tamás Zeke, Szabolcs Nováczki, Gábor Jeney: „Protocol Design and Analysis of a HIP-based Per-Application Mobility Management Platform", in the proceedings of the 7-th ACM International Symposium on Mobility Management and Wireless Access (MobiWAC 2009), ISBN:978-1-60558-617-5 , DOI: 10.1145/1641776.1641779, pp. 7-16, Tenerife, Canary Islands, Spain, 26-27 October 2009.

[C19] Gábor Jeney, László Bokor, Zsigmond Mihály: „GPS Aided Predictive Handover Management for Multihomed NEMO Configurations", in the proceedings of the 9-th IEEE International Conference on ITS Telecommunications (ITST'09), E-ISBN: 978-1-4244-5347-4, Print ISBN: 978-1-4244-5346-7, DOI: 10.1109/ITST.2009.5399380, pp. 69 – 73, Lille, France, 20-22 October 2009.

[C20] László Bokor, Zoltán Kanizsai, Gábor Jeney: „Performance Evaluation of Key IMS Operations over IPv6-capable 3G UMTS Networks", In Proceedings of 2010 Ninth International Conference on Networks, ICN'10, ISBN: 978-0-7695-3979-9, Print ISBN: 978-1-4244-6083-0, DOI:10.1109/ICN.2010.49, pp. 262 - 271,Menuires, France April 11-16, 2010. (Received Best Paper Award)

[C21] L. Bokor, Z. Faigl, S. Imre, „A Delegation-based HIP Signaling Scheme for the Ultra Flat Architecture", Proceedings of the 2nd International Workshop on Security and Communication Networks (IWSCN 2010), ISBN: 978-91-7063-303-4, pp. 9-16, Karlstad, Sweden, May 26-28, 2010.

[C22] Z. Faigl, L. Bokor, P. M. Neves, R. A. Pereira, K. Daoud, P. Herbelin, „Evaluation and Comparison of Signaling Protocol Alternatives for the Ultra Flat Architecture", Proceedings of The Fifth International Conference on Systems and Networks Communications (ICSNC 2010), ISBN: 978-0-7695-4145-7, Nice, France, Aug. 22-27, 2010.

[C23] R. Fracchia, C. Lamy-Bergot, G. Panza, J. Vehkaperä, E. Piri, T. Sutinen, M. Mazzotti, M. Chiani, S. Moretti, G. Jeney, L. Bokor, Z. Kanizsai and M. G. Martini, „System architecture for multimedia streaming optimisation", in Proc. of Future Network and MobileSummit 2010, Florence, June 2010.

[C24] Szabolcs Kustos, László Bokor, Gábor Jeney, „Testbed Evaluation of Dynamic GGSN Load Balancing for High Bitrate 3G/UMTS Networks", Proceedings of the IEEE 73rd Vehicular Technology Conference (VTC2011-Spring), pp. 1-5., ISBN: 978-1-4244-8332-7, DOI: 10.1109/VETECS.2011.5956406 , Budapest, Hungary, 05.15-05.18., 2011.

[C25] József Kovács, László Bokor, Gábor Jeney: „Performance Evaluation of GNSS Aided Predictive Multihomed NEMO Configurations", In: ITST-2011: 11th International Conference on ITS Telecommunications. Szentpétervár, Oroszország, Institute of Electrical & Electronics Engineers (IEEE), pp. 293-298.(ISBN: 978-1-61284-670-5), 2011.

[C26] Zoltán Faigl, Jani Pellikka, László Bokor, Sándor Imre, Andrei Gurtov: „HIP in 3GPP EPC", In: IETF 82 Proceedings. Taipei, Tajvan, 2011.11.13-2011.11.18. pp. 1-54. Paper HIPRG-4. , 2011.

[C27] Péter András Kara, László Bokor, Sándor Imre: „Distortions in QoE measurements of ubiquitous mobile video services caused by the preconceptions of test subjects", In: IEEE/IPSJ International Symposium on Applications and the Internet SAINT2012. Izmir, Turkey, 2012.07.16-2012.07.20. IEEE, pp. 409-413. ISBN: 978-0-7695-4737-4 , 2012.

[C28] Ivett Kulik, Péter András Kara, Tuan Anh Trinh, László Bokor, Analysis of the Relationship between Quality of Experience and Service Attributes for 3D Future Internet Multimedia, In: IEEE 4th International Conference on Cognitive Infocommunications. Budapest, Magyarország, 2013.12.02-2013.12.05. Budapest: pp. 641-646. ISBN: 978-1-4799-1544-6 , 2013.

[C29] Ivett Kulik, Péter András Kara, Tuan Anh Trinh, László Bokor, Attributes Unmasked: Investigation of Service Aspects in Subjective Evaluation of Wireless 3D Multimedia, In: The Second International Conference on Informatics & Applications (ICIA2013), Lodz, Lengyelország, 2013.09.23-2013.09.25. pp. 270-275. Paper 150. ISBN: 978-1-4673-5255-0, 2013.

[C30] László Bokor, Gianmarco Panza, Janne Vehkaperä, Lorenzo Iacobelli, Esa Piri, Matteo Mazzotti, Benoit Lecroart, Maria Martini, Cross-layer Optimized Delivery for Interactive Multimedia Healthcare Services: The CONCERTO Architecture, In: Future Network and MobileSummit 2013, Funems 2013. Lisboa, Portugália, 2013.07.09-2013.07.13. pp. 1-4. , 2013.

[C31] Péter András Kara, László Bokor, Sándor Imre, Distortions in QoE Assessment of 3D Multimedia Services on Multi-access Mobile Devices, In: 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). Lyon, Franciaország, 2013.10.07-2013.10.09. (IEEE) pp. 311-318., 2013.