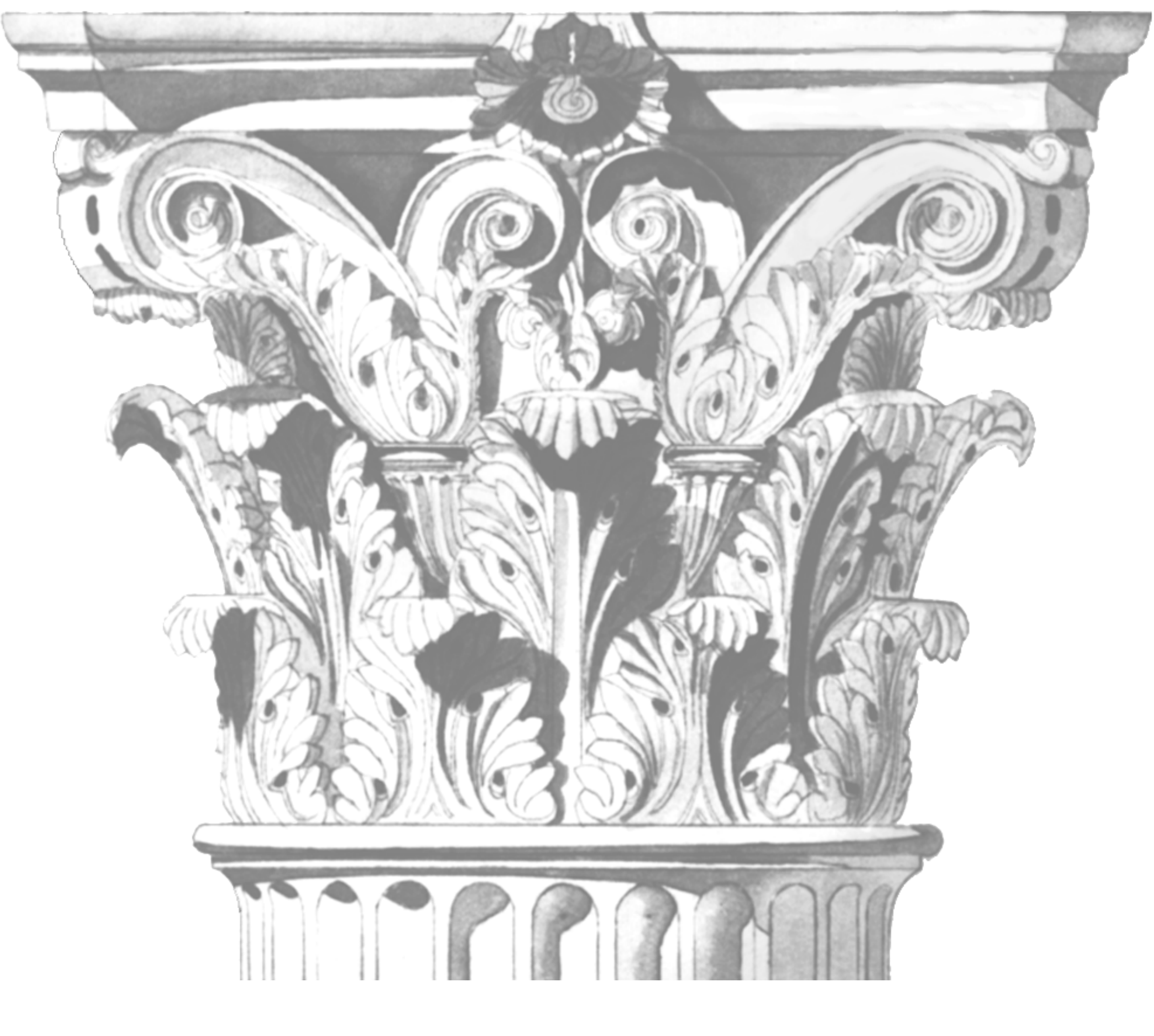
VERTICAL HANDOFF AND MOBILITY — SYSTEM ARCHITECTURE AND TRANSITION ANALYSIS

MIKA YLIANTTILA

Faculty of Technology,
Department of Electrical and
Information Engineering,
Infotech Oulu,
University of Oulu

OULU 2005



MIKA YLIANTTILA

VERTICAL HANDOFF AND MOBILITY — SYSTEM ARCHITECTURE AND TRANSITION ANALYSIS

Academic Dissertation to be presented with the assent of the Faculty of Technology, University of Oulu, for public discussion in Raahensali (Auditorium L10), Linnanmaa, on May 6th, 2005, at 12 noon.

OULUN YLIOPISTO, OULU 2005

Supervised by
Professor Kaveh Pahlavan
Professor Pentti Leppänen

Reviewed by
Professor Dharma Agrawal
Professor Kimmo Raatikainen

ISBN 951-42-7691-4 (nid.)
ISBN 951-42-7692-2 (PDF) http://herkules.oulu.fi/isbn9514276922/
ISSN 0355-3213 http://herkules.oulu.fi/issn03553213/

OULU UNIVERSITY PRESS
OULU 2005

Ylianttila, Mika, Vertical handoff and mobility — system architecture and transition analysis
Faculty of Technology, University of Oulu, P.O.Box 4000, FIN-90014 University of Oulu, Finland, Department of Electrical and Information Engineering, Infotech Oulu, University of Oulu, P.O.Box 4500, FIN-90014 University of Oulu, Finland
2005
Oulu, Finland

Abstract

The contemporary information age is equipped with rich and affordable telecommunication services. In the future, people have even more flexibility when true wireless Internet and real-time multimedia are provided seamlessly over heterogeneous wireless networks. Optimally combining the capacity and services of the current and emerging networks requires a holistic view of mobility, resource and service management. This thesis contributes to the research and development of these hybrid systems with three main contributions.

Firstly, a system architecture for vertical handoff in location-aware heterogeneous wireless networks is proposed. The proposed architecture enables the mobile node to prepare for approaching vertical handoffs and wake-up a hotspot interface. The needed communication procedures are discussed, and inter-related issues of mobility and geolocation information are considered in proportion to usability, advantages and limitations.

Secondly, a framework for the analysis of vertical handoff algorithm sensitivity to various mobility parameters including velocity, handoff delay and dwell time is introduced. Handoff smoothing with a dwell-timer is analyzed as one potential scheme for optimizing vertical handoff locally. It is compared to a power based algorithm to find out its sensitivity to the changes in effective data rates, velocity of the terminal and the amount of handoff delay. The analysis focuses on the transition region, having case studies on both moving-in and moving-out scenarios. An optimal value for dwell-timer is found through simulations, showing a performance gain over power based algorithm as a function of mean throughput. The analysis is extended also to a multiple network scenario.

Thirdly, experimental results on the behaviour of protocols used in wireless IP networks are presented. Prototype systems demonstrate results of using Mobile IP with a fuzzy logic algorithm for vertical handoff in a heterogeneous network environment and the role of IPv6 when using a voice application in a wireless LAN environment. Latest contributions include developing plug-and-play middleware functionalities for Symbian mobile devices, extending the use of the earlier results to state-of-the-art mobile devices.

Keywords: algorithms, heterogeneous wireless networks, interworking, location-awareness, mobility management, system architecture, transition analysis, vertical handoff, wireless local area network, wireless systems

Preface

"The best scientists are poets; the real engineer is an artist." – Sue Birchmore

In my attempt to prove my worthiness as a scientist and researcher, I start my thesis with a poem. While I wrote the poem mainly thinking about future researchers who plan their doctoral studies, it is also a summary of some of the lessons learned while preparing this thesis.

> *"If dissertation is your goal, you should know more than where is the North pole. Persistence and energy, you will need them as you will see. Do not start with haste and too high speed; the journey will teach you the lessons you need. You learn there is no gain without pain, whether it is in physical or intellectual plane. Completion is easier said than done, hard work must first come. Writing code, papers and reports; sometimes getting to conferences at nice resorts. Step by step you must go; through frustration and mental growth. Find a balance between work and rest; you can not do better than your best. Play it like a long game, that way you will remain (more or less) sane. There are times of trouble, but at the end you can say: 'I did it my way'. But, not without many thanks to say, to all the people who have helped me on the way. I say: thank you!"*

I can remember very vividly when I came to discuss about starting my diploma thesis work at the Centre for Wireless Communications (CWC) fall of 1997. CWC director at that time, Dr. Jaakko Talvitie asked me if I would be interested about a topic concerning interconnecting different wireless networks. I said yes with a little hesitation, and my journey started from there. Before starting my diploma thesis work, I made my first technical survey report for selecting a wireless network suitable for piloting virtual reality applications in the PIHVI project. After that I started working in the WiLU project in early 1998. That was my school for research, especially in the direction of Professor Kaveh Pahlavan who became my supervisor during the licentiate and doctoral thesis. Before that I have had the privilege to have Professor Savo Glisic supervising my diploma thesis work. During the WiLU project I went for the first time oversees to the USA, and visited Worcester Polytechnic Institute (WPI) several times at the Center for Wireless Information Network Studies (CWINS) at the Atwater Kent Laboratories of

WPI. I remember very well the meetings and technical discussions with the fellow researcher at WPI, including Prashant Krishnamurthy (currently professor at the University of Pittsburgh), Ahmad Hatami and many others. We started our work with technical implementations of Mobile IP based solutions on Linux and analytical studies of various handoff algorithms. That resulting learning process formed the basis for this thesis.

This research work would not be possible without long-term funding by the Finnish Technology Agency TEKES and industrial partners. Project partners in the various research projects include CWC, Oulu, Finland. CWINS, Worcester, USA; VTT, Oulu, Finland; Mediateam, University of Oulu, Finland; Nokia, TeliaSonera Finland, Finnish Defence Forces, Elektrobit, IBM and Serv-IT. Projects include PIHVI, WiLU, WINGIP, WINNER, LEMMINGS, FUTURA, MGAIN and most recently the Application Supernetworking (All-IP) project. They have had their own important role which can be seen in many places in this thesis. I would like to thank the fellow researchers, and steering and technical group members in these projects for their contributions and technical advice. This work has also been supported with scholarships by Tauno Tönningin Säätiö, Seppo Säynäjäkankaan Tiedesäätiö, HPY/Elisa Tutkimussäätiö, Nokia Oyj Säätiö and TeliaSonera Oyj tutkimus- ja koulutussäätiö.

For the supervision of this thesis I would like to express my special gratitude to Professor Kaveh Pahlavan. His contribution in directing and mentoring my research work has been conclusive. I have found him to be a man of wisdom, from whom I have learned a lot. I would like to also thank Professor Pentti Leppänen for co-directing this thesis, and Professors Petri Mähönen, Prashant Krishnamurthy and Jaakko Sauvola for their valuable comments and inspiring discussions during the years. For the pre-examination of my thesis and constructive comments improving the final outcome of this thesis I would like to thank Professor Dharma Agrawal from the University of Cincinnati, Ohio, USA, and Professor Kimmo Raatikainen from the University of Helsinki, Finland. Professor Abbas Jamalipour, who is the opponent in the defence together with Professor Raatikainen, deserves thanks for providing comments to an early version of the manuscript, affecting the title of the thesis. Zach Shelby provided high quality proof-reading comments.

From my fellow researchers in Finland I would like to especially acknowledge Juha-Pekka Mäkelä who has been the closest working colleague during this work. In addition, collaboration with Dr. Roman Pichna and Jari Vallström in the early days of CWC (both working currently at Nokia) helped to establish the foundation for all my future works. Finally, I would like to thank my wife Mari who is also preparing her doctoral thesis in the area of Biochemistry. Having her understanding (also as a researcher), empathy and love has provided me the energy and support needed to get this thesis work done. My father Seppo, mother Leena, sisters Mari and Riikka, relatives (especially grandmother Alma) and special friends have had their own important supporting role. I thank you all.

March 20th 2005, Oulu   Mika Ylianttila

List of original publications

This thesis is based on the following ten original papers (Appendices I-X) which are referred to in the text by Roman numerals:

I Ylianttila M, Mäkelä J & Pahlavan K (2005) Analysis of handoff in a location-aware vertical multi-access network. Elsevier J Comp Networks 2: 185-201.

II K. Pahlavan K, Krishnamurthy P, Hatami A, Ylianttila M, Mäkelä J, Pichna R & Vallström J (2000) Handoff in hybrid mobile data networks. IEEE Personal Communications Magazine 2: 34-47.

III Ylianttila M, Pichna R, Vallström J, Mäkelä J, Zahedi A, Krishnamurthy P & Pahlavan K (1999) Handoff procedure for heterogeneous wireless networks. Proc. IEEE Global Telecommunications Conference, Rio de Janeiro, Brazil, 5: 2783-2787.

IV Ylianttila M, Mäkelä J & Pahlavan K (2000) Geolocation information and inter-technology handoff. Proc. IEEE International Conference on Communications, New Orleans, Louisiana, USA, 3: 1573 -1577.

V Hatami A, Krishnamurthy P, Pahlavan K, Ylianttila M, Mäkelä J & Pichna R (1999) Analytical framework for handoff in non-homogeneous mobile data networks. Proc. 10th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Osaka, Japan, 1: 760-764.

VI Ylianttila M, Pande M, Mäkelä J & Mähönen P (2001) Optimization scheme for mobile users performing vertical handoffs between IEEE 802.11 and GPRS/EDGE networks. Proc. IEEE Global Telecommunications Conference, San Antonio, Texas, USA, 6: 3439 -3443.

VII Ylianttila M, Mäkelä J & Mähönen P (2002) Supporting resource allocation with vertical handoffs in multiple radio network environment. Proc. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Lisbon, Portugal, 1: 64-68.

VIII Ylianttila M, Mäkelä J, Krishnamurthy P & Pahlavan K (2000) Inter-technology mobility testbed. Proc. Finnish Wireless Communications Workshop, Oulu, Finland, 1: 38-43.

IX Ylianttila M, Rapeli M, Mäkelä J & Hautala T (2001) Comparative analysis of VoIPv4 and VoIPv6 in a bandwidth-limited wireless LAN testbed. Proc. IEEE International Conference on Communications, Helsinki, Finland, 3: 743 -747.

X Harjula E, Ylianttila M, Ala-Kurikka J, Riekki J & Sauvola J (2004) Plug-and-Play Application Platform: towards mobile peer-to-peer. Proc. 3rd International conference on Mobile and Ubiquitous multimedia, College Park, Maryland, USA, 1: 63-69.

Papers I-IV include aspects for system architecture design. Papers I and V-VII present an analytical simulation framework with the results and analysis of two handoff algorithms. Papers VIII and IX discuss early empirical results based on testbed measurements, while Paper X discusses a new concept that utilizes vertical mobility as part of an intelligent middleware providing holistic connectivity management and plug-and-play features.

In addition, the author has written in book chapters, magazine articles, conference papers and technical reports in versatile fields fields of academic research. These works range from core competence in both theoretical and empirical studies of mobility in heterogeneous wireless networks to wireless applications, geolocation systems, ad hoc networking, and technology trend analysis.

List of symbols and abbreviations

C_C	Control channels
C_N	Non-realtime traffic channels
C_H	Handoff cut-off priority channels
C_R	Real-time traffic channels
Δ_i	Vertical handoff delay for network i
d	Distance of the MH from the BS
ε	Geolocation error margin
f	Location updating frequency
$F_{i(k)}$	Factor K reducing effective throughput in network i
G	Gain in throughput
η_i	Throughput reduction co-efficient in network i
μ	Parameter for negative exponential distribution
N	Number of handoffs
Ω	Ratio of the effective data rates
P_T	BS transmit power
$p(v)$	State propability
r	Range of the hotspot
R_i	Nominal data rate in network i
R_k	Resource type
$s(d)$	Correlated log-normal shadow fading sample
s_i	Effective data rate in network i
s_T	Mean throughput during the transition region
s_{tot}	Total throughput during the transition region
$T_{D(i)}$	Dwell-timer for network i
T_i	Fragments of time spent in network i during the transition region
T_T	Transition region duration
v	Velocity of the MN
v_{zw}	State of a single cell in the vertical mobility model
W	Call or traffic type
χ	Tranceiver sensitivity threshold
Z	Platform type

3G	Third Generation
3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project 2
AAA	Authorization, Authentication and Accounting
ADSL	Asymmetric Digital Subscriber Line
AP	Access Point
API	Application Programming Interface
B3G	Beyond 3G
BB	Baseband
BCCH	Broadcast Control Channel
BER	Bit Error Rate
BLER	Block Error Rate
BPSK	Binary Phase Shift Keying
BRAN	Broadband Radio Access Networks
BS	Base Station
BSC	Base Station Controller
BSS	Base Station Subsystem
BSSGB	Base Station Subsystem GPRS Protocol
BTS	Base Tranceiver Station
BU	Binding Update
CAC	Channel Admission Control
CDMA	Code Division Multiple Access
CDPD	Cellular Digital Packet Data
CIP	Cellular IP
CIR	Carrier to Interference Ratio
CLDC	Connected Limited Device Configuration
CN	Correspondent Node
CGF	Charging Gateway Functionality
CPU	Central Processing Unit
CoA	Care-of Address
Codec	Implementation of a Coding-decoding algorithm
CORBA	Common Object Request Broker Architecture
CS	Channel-coding Scheme
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DAB	Digital Audio Broadcasting
DAD	Duplicate Address Detection
DE	Decision Engine
DHCP	Dynamic Host Configuration Protocol
DHT	Distributed Hash Table
Diff-Serv	Differentiated Services
DPSK	Differential Phase Shift Keying
DQPSK	Differential Quadrature Phase Shift Keying
DSMA/CD	Digital Sense Multiple Access with Collision Detection
DVB	Digital Video Broadcasting
EAP-SIM	Extensible Authentication Protocol – Subscriber Identity Module
ECS	Enhanced Channel-coding Scheme

EDGE	Enhanced Data rates for Global Evolution
EDR	Enhanced Data Rate
EHA	Entering Hotspot Algorithm
EIR	Equipment Identity Register
ETSI	European Technical Standards Institute
FA	Foreign Agent
FCC	Federal Communications Commission
FEC	Forward Error Coding
FL	Fuzzy Logic
Fr.Rel	Frame Relay
FUTURA	Future Radio Access
GGSN	Gateway GPRS Support Node
GMSK	Gaussian Minimum Shift Keying
GPN	Global Positioning Network
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile communication
GSN	GPRS Support Node
GTP	GPRS Tunneling Protocol
HA	Home Agent
HAWAII	Handoff-Aware Wireless Access Internet Infrastructure
Hcon	Holistic Connectivity Management
HDE	Handoff Decision Engine
HDSL	High-speed Digital Subscriber Line
HLR	Home Location Register
HIP	Host Identification Protocol
HIPERACCESS	High Performance Radio Access
HIPERLAN	High Performance Radio LAN
HIPERLINK	High Performance Radio Link
HMIP	Hierarchical Mobile IP
HN	Home Network
HO	Handoff (or handover)
HOD	Handoff Daemon
HOPOVER	HandOff Protocol for OVERlay networks
HSCSD	High Speed Circuit Switched Data
HRFWG	Home Radio Frequency Working Group
IAPP	Inter Access Point Protocol
ICMP	Internet Control Message Protocol
ICT	Information and Communication Technologies
IDMP	Intra-Domain Mobility Protocol
IEEE	Institute of Electrical and Electronics Engineers
IEEE 802.11a/b/g/e/i	Wireless LAN standards
IETF	Internet Engineering Task Force
Int-Serv	Integrated Services
IOCTL	Input Output Control

IP	Internet Protocol
IPDC	IP datacast
ISP	Internet Service Provider
ITHO	Inter-technology Handoff
ITU-R	International Telecommunication Union, Radiocommunication Sector
IWF	Interworking Function
IWMSC	Interworking MSC
J2ME	Java 2 Platform, Micro Edition
JXME	JXTA Micro Edition
JXTA	Set of open and generalized peer-to-peer protocols
LAN	Local Area Network
LBS	Location Based Service
LHA	Leaving Hotspot Algorithm
LLC	Logical Link layer Control
LPC	Linear Predictive Coding
LPN	Local Positioning Network
MAC	Medium Access Control
MAHO	Mobile Assisted Handoff
MBWA	Mobile Broadband Wireless Access
MD-IS	Mobile Data Intermediate System
MEHO	Mobile Executed (or controlled) Handoff
MGAIN	Mobile Entertainment Industry and Culture
MH	Mobile Host (same as MN)
MHF	Mobile Home Function
MIDP	Mobile Information Device Profile
MIMO	Multiple Input Multiple Output
MN	Mobile Node
MNLP	Mobile Node Location Protocol
MSC	Mobile Services switching Centre
mSCTP	mobile Stream Control Transmission Protocol
MG	Mobility Gateway
MSF	Mobile Serving Function
MT	Mobile Terminal
NAHO	Network Assisted Handoff
NAT	Network Address Translator
NEHO	Network Executed (or controlled) Handoff
NIC	Network Interface Card
NN	Neural Network
NTP	Network Time Protocol
OFDM	Orthogonal Frequency Division Multiplexing
OS	Operating System
OSI	Open System Interconnection
P2P	Peer-to-peer
PBCCH	Packet Broadcast Control Channel
PCM	Pulse Code Modulation

PCMCIA	Personal Computer Memory Card International Association
PDA	Personal Digital Assistant
PDN	Packet Data Network
PHY	Physical layer
PIHVI	Personal Virtual Services Based on Picocellular Networks
PLMN	Public Land Mobile Network
PnPAP	Plug-and-play Application Platform
PS	Proxy Server
PSTN	Public Service Telephone Network
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RA	Router Advertisement (Mobile IP/ICMP)
RA	Routing Area (GPRS)
RADIUS	Remote Authentication Dial-In User Service
RAN	Radio Access Network
RDF	Resource Description Framework
RFIC	Radio Frequency Integrated Circuit
RLC	Radio Link Control
RPE-LTP	Regular Pulse Excited Long Term Prediction
RS	Router Solicitation
RSL	Received Signal Level
RSVP	ReSerVation Protocol
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RTP	Real-time Transport Protocol
SCS	Secure Coordinate Server
SGSN	Serving GPRS Support Node
SER	Symbol Error Rate
SIMPLE	SIP for Instant Messaging and Presence Leveraging Extensions
SIG	Special Interest Group
SIP	Session Initiation Protocol
SiP	System-in-a-Package
SIR	Signal to Interference Ratio
SM	State Machine
SMIP	Seamless Mobile IP
SMLC	Serving Mobile Location Center
SM-SC	Short Message Service Centre
SMS-GMSC	SMS Gateway Mobile Switching Center
SNDCP	Sub Network Dependent Convergence Protocol
SNR	Signal to Noise Ratio
SoC	System-on-Chip
SOHYP	Sum of Hyperexponentials
SWAP	Shared Wireless Access Protocol
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access

TE	Terminal Equipment
TTL	Time To Live
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
UWB	Ultra-Wideband
VAP	Virtual Access Point
VDSL	Very high speed Digital Subscriber Line
VLR	Visiting Location Register
VoIP	Voice over IP
WCDMA	Wideband Code Division Multiple Access
WiLU	Wireless LAN for UMTS
WiMAX	Worldwide Interoperability for Microwave Access
WINGIP	Wireless Indoor Geolocation and IPv6 traffic analysis
WINNER	Wireless Inter-technology Networks with optimizEd data Rates
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network
WPAN-HR	WPAN High Rate (Wireless Firewire)
WPAN-LR	WPAN Low Rate
WMAN	Wireless Metropolitan Area Networks
WWAN	Wireless Wide Area Networks
WWW	World Wide Web
XCAP	XML Configuration Access Protocol
XMPP	Extensible Messaging and Presence Protocol

Contents

1 Introduction

"The wireless telegraph is not difficult to understand. The ordinary telegraph is like a very long cat. You pull the tail in New York, and it meows in Los Angeles. The wireless is the same, only without the cat." -Albert Einstein

The topic of this thesis "Vertical Handoff and Mobility – System Architecture and Transition Analysis" reflects a paradigm shift towards new generations of mobile networks where seamless mobility across heterogeneous networks and services is made possible. These new generations are referred to as 3G+, Beyond 3G (B3G) or 4G, indicating the fact that seamless mobility among heterogeneous wireless access networks and services needs further research and development beyond the current third generation networks. For clarity, term B3G is used in this thesis to represent these technologies.

As foresightfully and playfully illustrated in Albert Einstein's aphorism, customers may not need to know what wireless technology, base station, access point or router they are using at a given moment – they only need to experience seamless service. We may childlike think of the system as a cat that meows in Australia when the tail is pulled in Finland. Telecommunication engineers and researchers can be thought of as secret agents on television: they design the systems and let people know and worry about technical details only on a "need to know basis". When technical details about seamless mobility and handoff across heterogeneous networks are considered - they don't need to know! Transfer from one technology to another should be made easy and transparent.

This thesis contributes to the evolution of networking technology by making vertical mobility more understandable as a technical problem, and by helping to make it "invisible" or transparent for the average user. This thesis presents the main contributions of the author's research studies during the past seven years along with an overview of the field.

This first chapter discusses the research questions and provides problem the statement and motivation. The scope and the methodology are described, and an overview of the contributions is given. Finally, the outline of the thesis is introduced.

1.1 Problem statement and motivation

The vertical mobility paradigm deserves research attention from many perspectives. This thesis focuses on finding a solution to the vertical mobility problem in three categories. The first question is what kind of added value and changes to the system architecture does location-awareness bring when considered from the view point of vertical mobility management. Elaboration of this problem requires analysing the system architecture and related communication procedures. The second question is how to utilize the asymmetric data rates in overlapping heterogeneous wireless networks to the fullest. Elaboration of this problem requires developing an analytical framework for handoff analysis that considers both moving-in and moving-out scenarios. The objective is to find a method that will give better performance than a traditional handoff algorithm based on a received signal strength threshold. The third question is how to analyze vertical mobility and handoff in a packet-switched all-IP empirical context. This involves developing a middleware approach that on the one hand looks up in the protocol stack at the application layer requirements from applications such as Voice over IP (VoIP) or mobile file sharing, and on the other hand looks down in the protocol stack and sees the underlying wireless radio network resources. This type of holistic approach requires experimenting and prototyping with both advanced handoff algorithms, mobility management solutions, mobile middleware and wireless applications. This type of problem statements calls for a hybrid and cross-layer analytical approach for utilizing aggregated radio resource management and inter-domain packet-switched routing in a collaborative manner.

The motivation for addressing the given problem areas is to develop wireless packet-switched communications by facilitating a holistic approach for handoff and mobility management in heterogeneous wireless networks. The cross-layer approach involves taking into consideration on the one hand the handoff algorithms that are employed in the physical layer solutions in cellular networks and on the other hand mobility management solutions that are employed in the network layer or above. Lower layer parameters and context information such as location-awareness can be used as triggers in prepararing for the handoff decision. The whole problem area involves designing systems that can meet the requirements of managing the system capacity for various applications and mobility scenarios for a growing user population. In fact, there is an insatiable need for more wireless bandwidth, low packet loss rates and low latencies. The faster, the cheaper and more usable that wireless technologies evolve, the better. Applications and operating systems are getting more resource consuming as they are becoming more capable for supporting richer multimedia, games and other content. The most challenging task for system designers is to provide means to deliver time sensitive data such as voice or video over a heterogeneous network environment such as the Internet where limitations or bottlenecks for bandwidth often exists. For holistic system design, selections have to be done between different implementation strategies for how to utilize available alternative protocols, algorithms, middleware and other technical solutions. For the end-user, it would be preferable to hide the specifics of the implementation and configuration, resulting in a seamless experience when using heterogeneous networks, services and applications.

Emerging new networking technologies provide new possibilities for application and service developers. Enabling seamless or optimized vertical mobility is about introducing new resources for the Internet developers' community consisting of, e.g., middleware and networking protocol designers and application programmers. The more resources such as bandwidth or processing power available, the better algorithms and techniques the application designers can use to exploit the available resources to the fullest. Vertical mobility interconnects both existing and also any new wireless radio technologies, whether it is designed to be wide range, short range, indoors or outdoors. Main challenge is to implement applications and software platforms that take full advantage of the diversity of the wireless networks. This thesis helps to understand the key parameters and issues in this diversity to be considered in the system architectural design, implementation and standardization. As vertical mobility means mobility between non-homogeneous technologies, one has to consider the factors that diversify the wireless technologies. Most notable differences consider effective data rates and end-to-end delay, both instrumentally affecting the perceived service quality and application usage experience of the end-user. In addition, service providers have to understand what type of applications can be used in various vertical mobility scenarios. For this, the process of handoff (or handover) has to be elaborated, both for moving in and out of wireless hotspots.

1.2 Scope and methodology

As was in the problem statement, the scope and methodology of this thesis have also three categories. Firstly, architectural issues in integrated geo-telecom networks are discussed and a new concept of using geolocation information in vertical mobility is presented. This line of work includes identification of the theoretical aspects around this concept and building a framework for analysis and future work. Also, a new communication concept around application supernetworking is developed.

Secondly, vertical handoff efficiency is one important area for traffic optimization in the next generation wireless networks. It is essential to manage the delay and throughput in the vertical handoff process. The characteristics of delay and throughput in vertical handoff between non-homogeneous systems such as 802.11 WLAN and cellular networks (GPRS, EDGE or UMTS) are analyzed with simulations. This thesis has a specific emphasis on the handoff situation with the limited scope of a transition region where the received signal strength varies around the sensitivity of the WLAN receiver. Transition analysis provides information about the key parameters affecting algorithm performance measured as mean throughput perceived by a single user. Sensitivity to handoff delay and velocity is compared between threshold and dwell-timer based algoriths. An analytical model is developed also to include more than two overlapping heterogeneous wireless technologies. The validity of the results must be evaluated in proportion to the limitations of the used simulation model. A simulation model is always an approximation of a real system, expressing only a portion of the whole truth of the studied phenomenon. The performance estimates of the simulation results have, at least, a theoretical value. While simulation studies can be easily extended for areas that are difficult to measure in a real-life testbed environment, experimental studies complement the analytical holistic approach. It must be noted that the results presented in this thesis were achieved during several research projects and separate tasks, thus resulting in some fragmentation in the overall scope. On the other hand, the methodology provides a wide aspect for overall analysis.

Thirdly, the thesis develops implementation-oriented wireless packet-based communications in hybrid wireless networks. The objective of the empirical work is to understand the behaviour of selected protocols in a wireless environment prone to communication errors and undeterministic delays. For vertical mobility management WLAN is the primary overlay technology. It is important to understand the operational radio environment and performance characteristics at the edge of the WLAN cell. The performance of a wireless Voice over IP (VoIP) application is studied to get a view of delay sensitive real-time application behaviour at the edge of the WLAN where the nominal data rate is 1 Mbps due to a multirate algorithm. As the Internet Protocol (IP) is evolving to a new version (v6) that offers a bigger address space at the cost of a bigger header size, there has been some concern about how the increased header size affects real-time traffic such as VoIP, especially in bandwidth-limited wireless systems that do not support header compression. Comparisons are done in respect of used codecs and the IP protocol version, considering protocol overhead (header size) in last hop wireless link communications. Quantitative measures such as average delay, delay variation, jitter (defined as a square root of the standard deviation) are used as a foundation for the analysis. For validity of the results error sources were minimized and the synchronization error was monitored.

Hybrid usage of a mobility management protocol for future ubiquitous Internet and networking algorithms for wireless connectivity is considered. Mobile IP based vertical mobility management used with a fuzzy logic algorithm is evaluated and compared with a hysteresis based algorithm. Furthermore, developing plug-and-play middleware functionalities for Symbian mobile devices is explored, extending the use of the earlier results to state-of-the-art mobile devices.

1.3 Contribution of the thesis

This thesis contributes to the given reseach problem area by utilizing architectural design, developing a theoretical and simulation framework and providing a proof-of-concept implementation and experiemental results. A brief overview of the contributions is given below. Detailed contributions in each paper are elaborated in Chapter 3 where the summary of contributions in the original papers is given.

This thesis provides an overview of the key concepts in next generation wireless networks and identifies important issues in the emergence of these networks. The usage of geolocation information in a multi-layer and integrated geo-telecom wireless networks is addressed in Paper I. This work extends the general architecture for mobility among heterogeneous wireless networks where different alternatives for vertical mobility were first studied in Paper II. Also aspects related to using geolocation information in assisting mobility management between non-homogeneous networks and handover procedures are discussed in Paper III. Architectural issues in integrated geo-telecom networks are discussed and new concept of using geolocation information in vertical mobility was presented in Paper IV. In addition to architectural considerations and communication aspects, the work includes theoretical aspects around this concept and builds a framework for analysis. The architecture enables mobile nodes with multiple battery consuming network interface to let local area hotspot network interface cards “sleep” when not needed and activate them when the mobile node is approaching the hotspot base station coverage area.

Simulations and analysis are used to compare the performance of two handoff algorithms and their sensitivity to mobility parameters such as handoff delay and mobile terminal velocity. Handoff smoothing with a dwell-timer is discussed as one potential scheme for optimizing vertical handoff. While the dwell-timer is well-know from its use in handoff for homogeneous networks, it is needed to relate its usage in asymmetric non-homogeneous networks to key performance metrics and find locally optimal values. Results show how important it is to control the handoff delay. Managing this delay is essential for optimizing throughput in vertical handoff and also for maintaining optimal service levels in heterogeneous wireless networks. The usage of a dwell-timer is not a complete solution, but a partial solution when optimizing vertical handoff. It is essential to minimize the delay, and also consider other performance enhancing methods to be used with the dwell-timer. The simulation model is based on earlier work that was published in Paper V where the author of this thesis was a co-author. The author of this thesis has continued this work and established a more comprehensive framework for analysis, published in Papers I, VI and VII. The analytical framework was also extended to consider multiple radio network environments. Handoff profitableness was formulated and rules for decision making and analysis of the profitableness of the decision were given.

Results from the empirical work shows that IP layer mobility management provides an efficient way to interconnect heterogeneous packet oriented networks. Simple hysteresis based and more complex fuzzy logic based algorithms were implemented into the handoff daemon program. It was seen that the fuzzy logic based algorithm performs better in the form of reduced handoff rate and optimized WLAN range usage, as discussed in Paper VIII. The analysis of the effects of IPv6 on the performance of VoIP in emerging wireless networks is discussed in Paper I. A comparison between VoIPv4 and VoIPv6 was done in a bandwidth limited WLAN system. The idea was to replicate the bandwidth available at the very edge of the WLAN. Handoff was not considered in this paper, but the experimentation with a voice application helped to understand the performance and bandwidth requirements of wireless VoIP in proportion to selecting between IPv4 and IPv6. Comparisons were done with both varying ICMP payloads and with actual VoIP applications with three codecs. The analysis of protocol performance was done both numerically and experimentally.

As discussed in Paper I, Holistic Connectivity (HCon) management includes intelligent middleware that uses lightweight state machines for run-time parameterisation to provide rules for switching network connectivity and triggering the vertical handoff process when needed. The author of this thesis is a researcher and project manager in a new project aiming at innovating new lead applications, developing a reference framework and environment for creating and testing new mobile and peer-to-peer (P2P) application scenarios, and the management of mobility and connectivity, the availability and usability of services, and the interoperability and business aspects of novel supernetworked applications.

The author has worked as a researcher and project manager in various projects at the University of Oulu during 1997-2005 with contribution in the planning, implementation, and management of these projects in addition to actual research work.

1.4 Outline of the thesis

The outline of this thesis is as follows. In this chapter the importance of the studied topics for next generation networks was shortly discussed. This will be further enlightened in Chapter 2 which provides an overview of the vertical mobility technology evolution and background, ranging from the first experimental systems, and going through the current state-of-the-art technologies towards future B3G and 4G systems. The whole area is looked at from a system architecture design point of view to form a combined approach for holistic system architecture design including a comprehensive understanding of heterogeneous networks, integration models and mobility management with location-awareness. A high level view for the issues related to mobility engineering, and resource and service management is provided. Mobility management is elaborated through mobility scenarios, the vertical handoff procedure and resource allocation aspects, having emphasis on transition analysis. In Chapter 3, the summary of contributions in the original papers is given, and conclusions are given at the end in Chapter 4, including directions for future work.

2 Overview of vertical mobility

"It is impossible to travel faster than the speed of light, and certainly not desirable, as one's hat keeps blowing off." – Woody Allen

This chapter of the thesis gives an overview of handoff and mobility among heterogeneous networks – called vertical handoff and mobility. The term *vertical* refers to overlapping wireless networks and their hierarchical and asymmetric relationship. The problem of vertical mobility can be illustrated with a simple tale. Let us think of a relay team of a rabbit and a mouse. They have a task to carry a carrot as fast as they can. The mouse cannot carry very heavy loads, but it can go far. The rabbit, in this illustration, is restricted inside a fence. Inside a fence it can carry a heavy load of carrots, but at the fence it has to give the carrots to a mouse. In this thesis the rabbit represents a wireless local area network (WLAN), and the mouse a GPRS network. This work focuses on the event where the rabbit and mouse exchange the carrot load. This event is called *handoff* and in more precise terms *vertical handoff*. The corresponding process is referred as a *transition* from GPRS to WLAN and vice versa. In *mobility management*, the mouse and rabbit communicate with each other about how and when they plan to exchange the carrot, and they may inform others in the community about it too. The carrot represents the *payload* that the mouse (GPRS) and the rabbit (WLAN) are carrying. The payload means data such as bits or bytes, which are parts of bigger files (such as multimedia files) in a certain format. The problem is to make the best out of a rabbit and a mouse carrying a carrot together. There are physical and practical limits in achieving that, but it is needed to build a systematic approach in the analysis to find out what is possible, what is desirable – and what is not.

In the following discussion an overview is given for vertical mobility related technologies. The convergence of telecommunication and data networks has been driven by the integration of Internet technologies with wireless and cellular networks. This integration involves both radio carrier switching, network level routing information updating, and transport and application level adaptation. The primary goal has been to enable these processes to appear seamless to the users. In addition, services provided by the operators have to adapt to the vertical mobility scenarios, and in some cases provide additional value through content-aware solutions. Service quality while switching between dissimilar systems has to be managed in a systematic way. These issues are discussed in this chapter with references to the key technical papers in the literature.

2.1 Background: a short history of vertical mobility

During the past decade both telecommunication and Internet technologies have been in a phase of rapid development. Till the beginning of the new millennium the development was mainly technology driven and the real user needs were many times forgotten. Still, the mobile Internet evolution has taken many important steps towards providing better quality wireless data services to a wide audience. In cellular networks, evolution for the first three generations contributed to growing data rates and enhanced communication capabilities, achieving its current peak only recently in the third generation (3G) mobile networks and handsets. At the same time wireless local area networks have achieved enormous popularity in providing wireless broadband connection in public, enterprise and residential environments. Combining these two wireless technologies has attracted researchers now for about a decade, but there still remains issues to study. The next evolutionary steps after the third generation aim to provide extended mobility with optimized data rates and services. Nomadic users have more flexibility when using multi-service networks that provide services such as seamless connection to the Internet via heterogeneous networks, advanced spatial location and navigation services and true IP based real-time multimedia. One of the key challenges in future network management is end-to-end optimization that takes into account variables such as throughput optimization, routing optimization, delay profiles for heterogeneous wireless environments and also economical profitability. The door for next generation networks and services beyond 3G is opening and is soon ready for entering. These systems are called B3G (beyond 3G) or 4G. They will make heavy use of heterogeneous networking technologies.

2.1.1 Evolution of personal communications

The use of Internet services is done already over heterogeneous networking technologies. These technologies have gone though an evolutionary process. I remember very well my brand-new Pentium 120 MHz in 1996. Back then, just using the Internet from home was a thrill. Being able to read e-mail, update a personal web-site and view other people’s home pages was really all I needed. A 14.4 kbps modem connection did not feel at all too slow, nor the 120 MHz processor. Actually, the 120 MHz Pentium processor was pretty much the fastest the money (or a student) could by at the time. But since then, the requirements for “fast” have increased every year following Moore’s law which states that the processing capacity doubles every 18 months or so. Soon there were 300 MHz processors, then 800 MHz, and rising over 1 GHz. Modem speeds increased through 28.8 kbps to 56 kbps variable bit rate capacity. Now I have at my home a 2 GHz processor and 2/2 Mbps VDSL broadband connection. I am able to share that bandwidth with, e.g., a laptop attached to my TV set. Thus I can use my home PC as a repository for movies and music files, which can be shared over wired Ethernet or wireless local area network connection at my home.

The use of “mobile Internet” was started in the early 1990’s with GSM data connections offering a nominal data rate of 9.6 kbps (in practice much less than that). It was both very slow and expensive when compared to wired use of the Internet. Then, in the latter half of the 1990’s GPRS was introduced offering up to 40 kbps nominal data rates, depending on how many time slots were allocated to use. Yet, GPRS suffers from high peak delays and delay variations. Currently, higher data rate services are being made possible with EDGE and UMTS services. These services still tend to be expensive, especially in comparison to WLAN. Already in the late 1990’s the use of WLANs grew in popularity as an extension to a fast broadband connection in office environments. In the beginning of the new millennium they started to become more common in public settings such as airports and hotels, and marching their way also to home environments. This happened partly due to broadband xDSL technologies becoming more common.

Now I can use at home Internet over a broadband wireless connection, and on the road I can use a laptop with a GPRS, EDGE or UMTS connection. Thus I can be “on-line” whenever I want, thus fulfilling the original aims of mobile Internet: anywhere, any time. I have a monthly rate for my GPRS connection, being able to download 100 Mbytes of data with an affordable constant monthly fee that is about 20 €. Faster connection speeds with UMTS are becoming available, but the price for that is at the moment not very affordable considering heavy Internet use. While currently operators in Finland promise 64-384 kbps connection speeds with EDGE and UMTS with an affordable price about 20 € when 100 Mbytes of data is transferred in a month, the price gets higher when that 100 Mbytes is exceeded. If regular customers needs to transfer hundreds of Mbytes, wireless cellular data is currently not a very cost-effective solution. Using a wireless or wired connection linked to a home broadband connection is much more cost-efficient. This is one reason to consider interconnected use of heterogeneous wireless connections: heavy use of the Internet (downloading a large amount of data) can be postponed to home premises, while it may be needed to maintain Internet connectivity when transferring from “on-the-road” to the home domain. For example, one may want to keep alive connectivity to some databases without the need to re-establish the connection when changing the domain of operation. Seamless transfer of connectivity is thus need over heterogeneous networks to enable such a scenario.

Optimized connectivity management over heterogeneous networks is driven by all-IP applications that set a continuously growing need for more wireless bandwidth. A mobile user should be able to use them over the TCP/IP protocol suite which has established its role as the *de-facto* communication protocols standard family. All-IP applications include all the standard Internet applications such as WWW, E-mail, file transfer and text chat applications. In addition, peer-to-peer (P2P), online games, virtual reality applications, VoIP (Voice over IP), video conferencing and instant messaging applications are getting more popular. In P2P, a mobile user may, e.g., share files from his/her computer over the Internet. Application supernetworking combines different features of various applications in a unified session. As illustrated in Fig. 1, these applications and their wide use in the Internet set increasing requirements for wireless access network bandwidth and practical data rates (throughput). In fixed networks, applications such as VoIP have been a target of public and industrial interest. An example of a recent VoIP application is Skype[1], which uses a P2P protocol in establishment of connections. It also provides a gateway to any telephone number. One can make a VoIP call from a laptop to a mobile phone, for instance. There are also many teleoperators and ISPs proving long distance VoIP services with a lower rate than the legacy PSNT alternatives. A logical step is to utilize VoIP also in wireless networks, but that requires more wireless capacity from the wireless radio access network (RAN). Considering the current pricing models of wireless cellular data, it may be that the most reasonable wireless VoIP scenario is WLAN or WPAN local wireless connectivity attached to a fixed broadband connection in home and office environments.

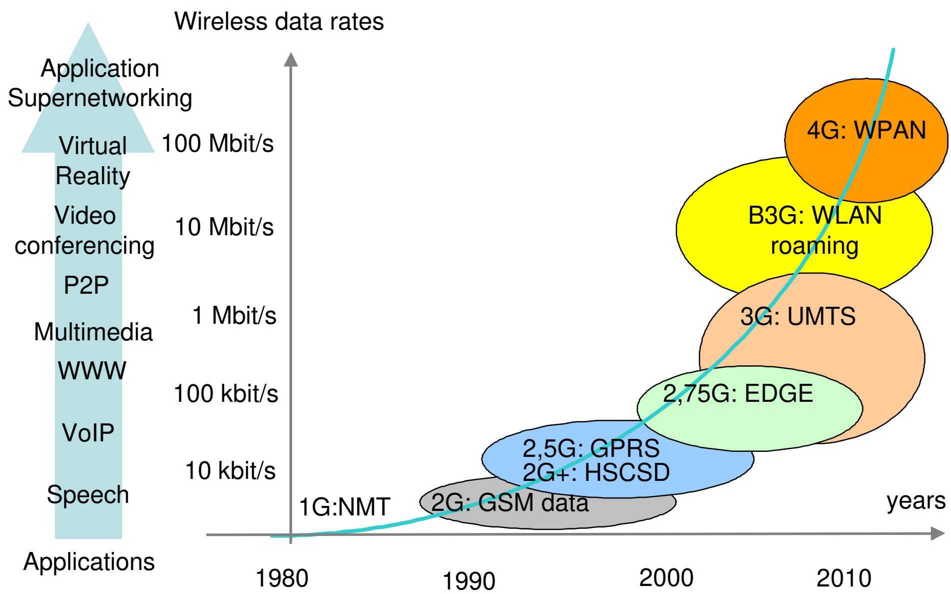


Fig. 1. Applications drive the need for higher wireless data rates.

Peer-to-peer (P2P) networks evolve from the existing third generation by providing mobile end-to-end services. Advanced mobile services develop through rich calls and multisessions to supersessions. In a rich call, a legacy phone call is enriched with additional video, picture and data content in the same session. In a multisession, a subscriber can communicate with one end-point using a voice session, and with another using a video session simultaneously. A supersession is a new concept that combines the features of a rich call and a multisession. Terminals with different capabilities can communicate with each other using versatile session management. The concept of application supernetworking ranges from simple file sharing applications utilizing superdistribution to complex applications that use supersessions to combine media streams from various peers. The efficient implementation of mobile applications and supernetworking in P2P networks requires an intelligent middleware in the mobile

[1] Voice over Internet application Skype www-page (accessed 19.10.2004) http://www.skype.com

handsets to take full advantage of the available network connections. Applications can also be made aware of the context (e.g., location) by integrating different services and technologies on a common platform.

2.1.2 Interconnecting heterogeneous networks

For scientists the existing diversity of both wired and wireless technologies presents two obvious questions: can we combine the use of these and what are the benefits of doing so? Integration, convergence and interoperability had been an important part in other technology evolutions, why not here. Inter-system mobility expands the possibilities of traditional communication networks by enabling more bandwidth or enriched services for mobile users. Especially, roaming between WLAN and UMTS (and its predecessors GPRS and EDGE) has been seen as an important scenario for both enabling new services and as a tool for the operator to balance network load. The research and development of this topic has been active since the early 1990's.

Since the early 1990's there have been numerous industrial initiatives and academic projects either directly or indirectly working with vertical mobility. The term *vertical mobility* was first introduced by the Daedalus/BARWAN[2][1] project at University of California Berkeley. Fig. 2 shows the conceptual architecture of vertically overlapping heterogeneous wireless networks, illustrating also the various domains of operation for the subnetworks in the overall architecture. Development around providing mobility to the famous Internet protocol suite TCP/IP started in the early 1990's [2]. The Internet Engineering Task Force (IETF) established a work group[3] for providing mobility support at first for the basic IPv4 protocol, and then later to the new IPv6 protocol. The academic networking community started to get interested at the same times about providing mobility across heterogeneous networks.

Network layer solutions using Mobile IP [3] attracted many academic projects as it provided a quick-start to demonstrate and experiment with a real-life testbed. Testbed setups in most cases consisted of a set of laptops equipped with wireless and wired network interface cards running Linux or FreeBSD operating systems. By adding a router or gateway to access the public Internet, and by designing and configuring the local test network, the setup was ready. The early academic releases of Mobile IP provided in the Internet with open source codes were MosquitoNET[4] [4], an implementation of Mobile IPv4 on the Linux OS at Stanford University, and Monarch[5],an implementation of both Mobile IPv4 and Mobile IPv6 on the BSD OS at Carnegie Mellon University.

---

[2] Bay Area Research Wireless Access Network www-page (accessed 19.10.2004) http://http.cs.berkeley.edu/~randy/Daedalus/BARWAN/BARWAN_index.html

[3] IETF work group for IP Routing for Wireless/Mobile Hosts WWW-site (accessed 19.10.2004) http://www.ietf.org/html.charters/mobileip-charter.html

[4] The Stanford University MosquitoNet project WWW-site (accessed 19.10.2004) http://mosquitonet.stanford.edu/

[5] The Rice University Monarch project WWW-site (accessed 19.10.2004) http://www.monarch.cs.cmu.edu/

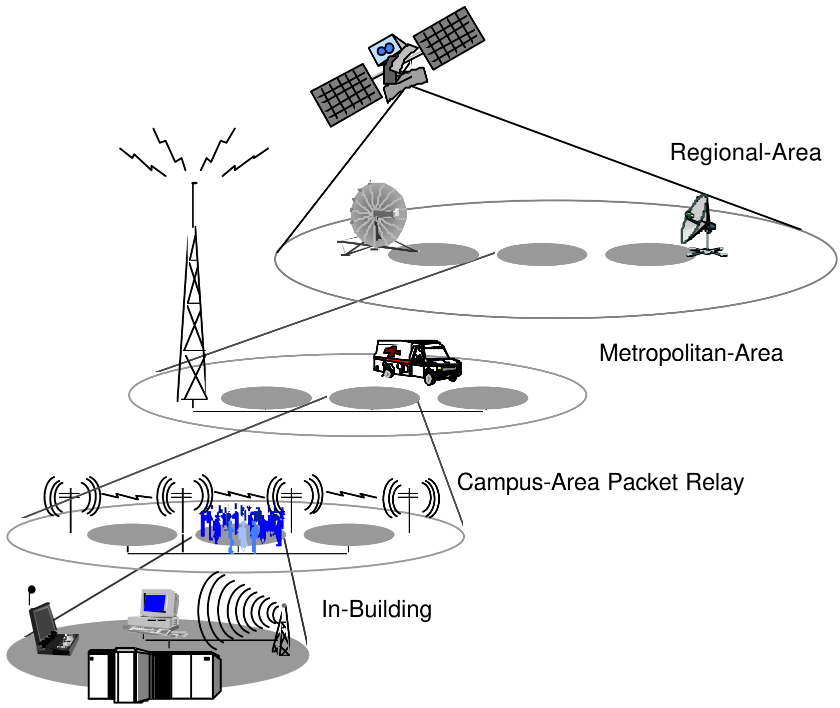


Fig. 2. Wireless overlaid heterogeneous network architecture [1].

The development around the Linux based testbed environment became the most famous form, following the general increase of the Linux OS in popularity. Mobile IPv4 implementations for Linux include those from the University of Singapore[6], State University of New York at Binghamton[7] and Dynamics[8] at the Helsinki University of Technology (HUT). Mobile IP performance with vertical mobility has been evaluated in projects such as Internode[9]. Mobile IPv6 implementations for Linux include that from Lancaster[10] University (UK) and MIPL[11] Mobile IPv6 at HUT. FreeBSD projects such as

---

[6] University of Singapore Mobile IP WWW-site (accessed 4.10.2004) http://mip.ee.nus.edu.sg/

[7] State University of New York at Binghamton Mobile IP WWW-site (accessed 4.10.2004) http://anchor.cs.binghamton.edu/~mobileip

[8] The Dynamics Mobile IP (originally developed at Helsinki University of Technology, HUT) WWW-site (accessed 19.10.2004) http://dynamics.sourceforge.net/

[9] EU project "Interworking of Nomadic Multi-Domain Services." IST-l999-20117, IST 5th Framework Programme R&D project, 2002.

[10] The Lancaster University Mobile IPv6 WWW-site (accessed 19.10.2004) http://www.cs-ipv6.lancs.ac.uk/ipv6/MobileIP/

[11] Mobile IPv6 for Linux project at Helsinki University of Technology, HUT WWW-site (accessed 19.10.2004) http://www.mipl.mediapoli.com/

the Inria Hierarchical Mobile IPv6 Proposal[12] provided source code for testing new initiatives for a new hierarchical structure in mobility routing. EU projects such as 6Net[13] have made extensive studies on the performance of Mobile IPv6.

The University of Oulu have studied Mobile IP performance in projects such as WiLU and WINNER by focusing on the handoff process, demonstrated using various handoff algorithms with Mobile IP. As an example, Paper VIII of this thesis discusses the use of a Fuzzy Logic (FL) algorithm with Mobile IP. This early work was one of the first contributions that combined a telecommunication oriented research related handoff algorithms such as FL or dwell-timer with computer science oriented research focusing on packet routing. This is one of the key areas of this thesis: to combine the aspects of using various handoff methods in the packet based all-IP communication environment. Especially, discussing the metrics for triggering handoff and finding the optimal dwell-timer value to optimize the handoff process at the edge of the wireless LAN cell is considered in this thesis.

There are many ways to utilize the vertical mobility paradigm. For example, the Application Supernetworking project[14] considers vertical mobility as a part of intelligent middleware called the Plug-and-Play Application Platform (PnPAP) that enables seamlessly using different communication protocols (such as SIP or XMPP) and connectivities (such as GPRS, EDGE or WLAN). In addition to academic projects, there are currently many commercial products available for providing mobility support across heterogeneous IP networks. These products include software such as Netseal[15] and Secgo[16] that provide automatic selection between available network connections added with security measures. More notably, in upcoming 3G devices Mobile IP is supported in the standards made by standardization bodies such as 3GPP[17] and 3GPP2[18].

As a technology, vertical mobility is still finding its way to the mass markets. Projects such as VHO[19] aim at increasing understanding of how multi-access will affect the services and end users. Field tests and usability studies are done also in other projects, such as Rotuaari[20] and COMS[21]. There are probably many other projects that are directly or indirectly related to vertical mobility, but the purpose here is to mention only some

---

[12] INRIA Hierarchical Mobile IPv6 Proposal (HMIPv6) WWW-site (accessed 19.10.2004) http://www.inrialpes.fr/planete/people/bellier/hmip.html

[13] European Union funded project 6Net deliverables WWW-site (accessed 19.10.2004) http://www.6net.org/publications/deliverables/

[14] AII-IP Applications Creation, Networking and Service Management project WWW-site (accessed 19.10.2004) http://www.mediateam.oulu.fi/projects/allip/

[15] Netseal company WWW-site (accessed 19.10.2004) http://www.netseal.com/

[16] Secgo company WWW-site (accessed 19.10.2004) http://www.secgo.com/

[17] The 3rd Generation Partnership Project (3GPP) WWW-site (accessed 19.10.2004) http://www.3gpp.org/

[18] The 3rd Generation Partnership Project 2 (3GPP2) WWW-site (accessed 19.10.2004) http://www.3gpp2.org/

[19] VHO project WWW-site (accessed 19.10.2004) http://www.cs.hut.fi/~pmrg/VHO.html

[20] University of Oulu Rotuaari project WWW-site (accessed 19.10.2004) http://www.rotuaari.net

[21] University of Cambridge Open Mobile System proecjt WWW-site (accessed 19.10.2004) http://www.cl.cam.ac.uk/Research/SRG/netos/coms/

case examples of the different types of vertical mobility related projects. There can be seen a global evolutionary progress towards eventual deployment of vertical mobility services with global roaming possibilities.

2.2 System architecture and design issues

The evolution and convergence of the cellular and computer industries has opened the road for many new and exciting applications and services. There is an obvious trend towards the convergence, integration and interoperability of various networking related technologies. In the wireless domain, the existing macro-, micro- and picocell networks often have overlapping areas of coverage. In next generation wireless systems called B3G or 4G the mobile users should be able to move among these heterogeneous networks in a seamless manner. Network operators and service providers must consider robust technological and architectural design aspects in enabling new multihomed services and applications. The enabling technological building blocks are discussed in the following.

2.2.1 Features of B3G systems

The first visions of B3G or 4G estimate this next generation to be realized around 2010 and based around five elements: fully converged services, ubiquitous mobile access, diverse user devices, autonomous networks and software dependency [5]. In this vision, the seamless connection of heterogeneous networks include cellular data, 3G, WLAN/HIPERLAN, short range PAN/LAN/MAN and broadcast DVB/DAB services integrated through an IP based core network. Implementing this type of integrated system invokes many challenges in mobile handset design, wireless system discovery, terminal mobility, topological fault tolerance and survivability [6]. Realization of the envisioned B3G or 4G systems and technologies requires unified efforts in the areas of standardization and development of the enabling technologies, providing access network convergence in an evolutionary manner, designing system architectures with considerations on application, protocols, location, authentication and security [7]. The fragmented nature of the heterogeneous radio spectrum calls for new software defined reconfigurable radio concepts utilizing economical smart and adaptive antennas in MIMO radio channels, multi-standard (3G/4G/802.11/GPS/Bluetooth) transceiver architectures, wideband OFDM/multi-carrier modulation, and other advanced solutions implemented in RFIC [8]. They will provide new performance in achieving data rates up to 100 Mbps and support for streaming, multicasting, and downloads of 5-20Mbps even when travelling at 200 km/hr, over an IP-centric network. Terminal implementation technologies include system-on-chip (SoC) and system-in-a-package (SiP) technologies with reconfigurable building blocks [9].

From the service aspect one has to find a balance between public and private service domains by adapting multiple standards and service environments (home, office, outdoors, indoors) across multiple operators and service provider domains with ensured QoS, data privacy and information integrity, and taking into account user profile and terminal characteristics [10]. This requires a service architecture that supports integrated mobility management for heterogeneous wireless network environment, session control and mobility, AAA (Authorization, Authentication and Accounting) functionality, and profile based personalizable service management [11]. In order to achieve end-to-end QoS, the system architecture has to consider the handoff process in particular. During the handoff between heterogeneous systems the mobile user may experience significant variation is QoS due to handoff delay caused by message exchanges, multiple database accesses, and negotiation-renegotiation processes affecting the performance of both upper-layer protocols and applications [12]. This type of problema requires considerations at many layers of the communications: priority-based routing at the network layer, delay budget calculations and the optimization of individual parameters (e.g., the dwell-timer value in a handoff algorithm) in both locally (for a single user) and system-wide (for all users in a cell or a subnet).

A heterogeneous and fragmented cell architecture enables advanced scenarios for multihop and ad hoc type communications [13], although they present more challenges for QoS guarantees due to the nature of relayed communications. Location-awareness may bring a useful additional parameter for adaptive applications and routing methods. Adding location awareness to the system architecture and vertical mobility management is considered in Paper I.

2.2.2 Technical aspects of vertical mobility

Fig. 3 shows a model representing a taxonomy of the technical aspects related vertical mobility. Looking from a top-down perspective, one can distinguish between resource management, mobility engineering and service management categories in the system architecture design.

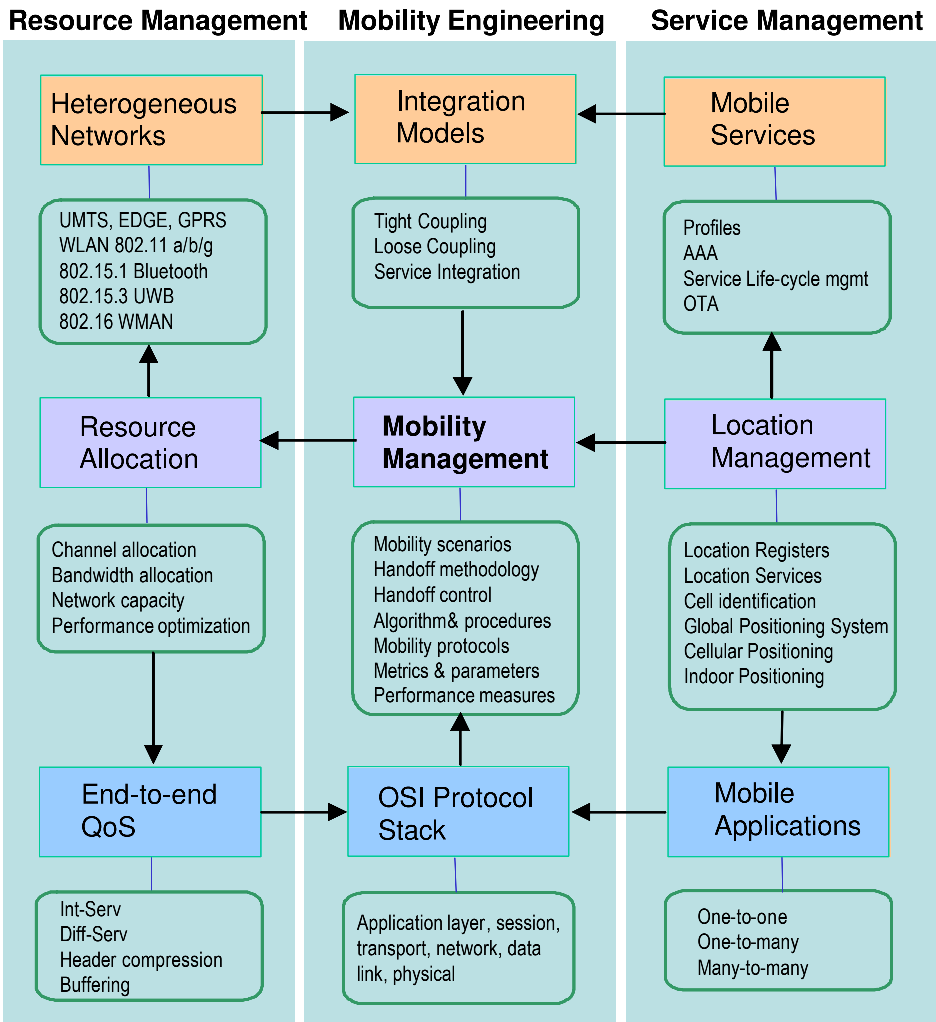


Fig. 3. System architecture design issues in vertical mobility.

Resource management is comprised of both *direct* (channel and bandwidth allocation,) and *indirect* (network capacity and performance optimization through various ways) resource allocation in a multiple heterogeneous wireless networks environment. Resource allocation affects directly the experienced QoS, but the end-to-end QoS also requires other management, such as prioritizing packets in the routing, using header compression over wireless links and buffering packets in the terminals and routers (limited by the QoS requirements).

Mobility engineering comprises integrating heterogeneous access networks and services, providing mobility management, and designing and implementing various protocols and middleware solutions in the different layers of the OSI protocol stack. At the core of the whole system architecture design is mobility management. As elaborated in Chapter 2.3, mobility management in the next generation all-IP based wireless systems uses hierarchical network architectures [14], both in horizontal and vertical planes. Mobility management across B3G networks, where "access is the killer application" [15], can be conceptually divided into several subtopics such as mobility and interworking scenarios [16], handoff control strategies, handoff algorithms, handoff performance measures, handoff methodology, handoff metrics and mobility parameters.

Service management includes providing mobile services (user profiles, AAA functions, and service life-cycle management through OTA functions (for downloading and upgrading services), location management (location registers and location services) and provisioning mobile applications.

This thesis focuses on mobility engineering through architectural design and performance estimations through simulations and measurements, having relations to location and resource management. The author's contribution of Papers I-X are summarized in Chapter 3. They are related to the design of the system architecture; including aspects of the overall system architecture along with an analysis of integrating location awareness to the architecture; mobility management, including analysis based on both vertical handoff algorithm simulation results, early empirical studies with testbed measurements and implementation aspects; and future directions.

2.2.3 Heterogeneous access networks

There have been a wide variety of wireless access network technologies developed during the past decade. As described in Table 1, wireless data access can be provided over heterogeneous wireless networks [17]. Some of the wireless network standards have aimed at specific market regions such as North-America, Europe and Asia, especially in cellular networks. In local and personal area networks there are some radio transmission regulations that differ between different regions, but there is more global competition in that area. In general, heterogeneous wireless access networks can be divided into wireless local area networks (WLAN), wireless personal area networks (WPAN), wireless metropolitan area networks (WMAN) and wireless wide area networks (WWAN), including cellular, satellite [18] and envisioned stratospheric [19] networks.

Table 1. Wireless data access in heterogeneous wireless network standards.

Network	Standard	Data rate (PHY)	Frequency band
Cellular networks	GSM data (2G)	9,6 kbps	900/1800/1900 MHz
	HSCSD (2G+)	14-42 kbps	“
	GPRS (2.5G)	14-128 kbps	“
	EDGE (2.75G)	128-384 kbps	“
	UMTS (3G)	up to 2 Mbps	1900 – 2025 MHz
WLAN	IEEE 802.11b	1, 2, 5.5., 11 Mbps	2,4 GHz
	IEEE 802.11a	1-54 Mbps	5GHz
	IEEE 802.11g	“	2,4 GHz
	IEEE 802.11e	“	2,4 GHz
	IEEE 802.11n	100 -540 Mbps	2,4 GHz, 5GHz
Bluetooth	IEEE 802.15.1	721kbps (BT 1.1) 2-20 Mbps (BT 2.0)	2,4 GHz
WPAN	IEEE 802.15.3	11-55 Mbps	2,4 GHz
UWB WPAN-HR	IEEE 802.15.3a	110 Mbps- 1 Gbps	3.1 – 10.6 GHz (FCC)
Zigbee WPAN-LR	IEEE 802.15.4	20-250kbps	868 MHz , 915 MHz 2,4 GHz
WMAN	IEEE 802.16a	75 Mbps	2-11 GHz
WiMAX	IEEE 802.16c	134 Mbps	10-66 GHz
WWAN MBWA	IEEE 802.20	2.25 – 18 Mbps	< 3,5 GHz
HomeRF	HRFWG SWAP	800 kbps – 10 Mbps	2,4 GHz
HIPERLAN	ETSI BRAN HIPERLAN/1	23,5 Mbps	5GHz
	ETSI BRAN HIPERLAN/2	1-54 Mbps	5GHz
	ETSI BRAN HIPERLAN/3 (HIPERACCESS)	25-100 Mbps	40.5-43.5 GHz
	ETSI BRAN HIPERLAN/4 (HIPERLINK)	up to 155 Mbps (150 m range)	17 GHz

In the following, cellular and wireless local area networking technologies are shortly discussed as base technologies for vertical mobility. Especially, vertical mobility between cellular technologies such as GPRS, EDGE and UMTS, and wireless local area networks such as IEEE 802.11b are considered in this thesis. Furthermore, the analysis can be extended to any combination of heterogeneous networks, including multiple overlapping (more than 2) radio technologies, as discussed in Paper VII.

Cellular networks in Europe have evolved through 2G (GSM data), 2G+ (HSCSD [20]), 2.5G (GPRS), 2.75G (EDGE) to 3G (UMTS/WCDMA), providing data rates from the 9.6 kbps of GSM data up to the 2 Mbps theoretical rate of UMTS. So far the dominating packet radio service over the cellular GSM core network in Europe has been GPRS (General Packet Radio Service). GPRS [21] uses exactly the same physical radio channels as GSM and only new logical GPRS radio channels are defined. The allocation of these channels is flexible: from 1 to 8 radio interface timeslots can be allocated per TDMA frame. Timeslots are shared by the active users, and the uplink and downlink are allocated separately. Physical channels are taken from the common pool of available channels in the cell.

Table 2. GPRS and EDGE transmit rates (PHY) for various channel coding schemes.

Channel Coding Scheme/ Transmit Rate (kbps)	*GPRS CS-1*	*GPRS CS-2*	*GPRS CS-3*	*GPRS CS-3*	*EDGE ECS-1*	*EDGE ECS-2*	*EDGE ECS-3*	*EDGE ECS-4*
1 timeslot	9.05	13.4	15.6	21.4	33.0	41.0	48.0	65.2
2 timeslots	18.1	26.8	31.2	42.8	66.0	82.0	96.0	130.2
3 timeslots (unidirectional max. of the first terminals)	27.2	40.2	46.8	64.2	99.0	123.0	144.0	195.6
4 timeslots	36.2	53.6	62.4	85.6	132.0	164.0	192.0	260.8
5 timeslots	45.2	67.0	78.0	107.0	165.0	205.0	240.0	326.0
6 timeslots	54.3	80.4	93.6	128.4	198.0	246.0	288.0	391.2
7 timeslots (1 for signalling)	63.4	93.8	109.2	149.8	231.0	287.0	336.0	456.4
8 timeslots (theoretical max. of the carrier capacity)	72.4	107.2	124.8	171.2	264.0	328.0	384.0	521.6
Modulation	GMSK	GMSK	GMSK	GMSK	O16-QAM	O16-QAM	O16-QAM	O16-QAM

Allocation to circuit switched services and GPRS is done dynamically according to a "capacity on demand" principle. This means that the capacity allocation for GPRS is based on the actual need for packet transfers. GPRS does not require permanently allocated physical channels. Logical network nodes called GPRS support nodes are used for packet routing in the backbone. The Gateway GSN (GGSN) acts as the interface to public data networks like the Internet and contains the routing information to be used to tunnel packets to the MH through a Serving GSN (SGSN). The SGSN is responsible for location management and the delivery of packets. The cellular data connections have developed rapidly from low data rates into high bit rate connections by, e.g., providing multi-slot connections with enhanced channel coding schemes. Table 2 shows the different modes of operation for GPRS and EDGE. In EDGE, the channel coding rates are enhanced with higher capacity modulation and coding schemes. However, the practical performance of these cellular data systems is usually much lower than the nominal rates, even down to ten times lower than the nominal rate [22].

As shown in Table 1, wireless local and personal area networking standards include IEEE 802.11 (WLAN), IEEE 802.15 (Bluetooth), IEEE 802.15.3 (WPAN), IEEE 802.15.4 (Zigbee), ETSI BRAN HIPERLAN and HomeRF SWAP has emerged aiming at short range wireless communication. In addition, IEEE 802.16 (WiMAX) and IEEE 802.20 (WWAN) aim at higher data rates over longer distances. In the wireless local area networks (WLAN), the IEEE 802.11 family of standards [23] has been most successful and so far has dominated the markets in this area.

Table 3. Range and PHY data rates for 802.11b WLAN.

Transmits Rate/ Range	"High speed" 11 Mbps	"Medium speed" 5.5 Mbps	"Standard speed" 2 Mbps	"Low speed" 1 Mbps
Open office environment	160 m	270 m	400 m	550 m
Semi-open office environment	50 m	70 m	90 m	115 m
Closed office environment	25 m	35 m	40 m	50 m
Receiver sensitivity	-83 dBm	-87 dBm	-91 dBm	-94 dBm

Table 3 shows a rough estimate about the modes of operation of IEEE 802.11b, which has been a dominating standard for many years. One must note that the practical rata rates observed by the user are noticeably lower due to various factors degrading the nominal throughput. Random access systems based on CSMA (Carrier Sense Multiple Access) have been found to be quite inefficient in cases where multiple users and real-time traffic is present, and performance can degrade down to 10-30 % of the nominal values, especially in the case of real-time applications used in a multi-user environment [24]. Performance is enhanced in IEEE 802.11g and IEEE 802.11a standards that employ OFDM (Orthogonal Frequency Division Multiplexing) technology. OFDM transmits multiple signals simultaneously over a single transmission path by distributing the data over a large number of carriers. These unique frequency ranges are spaced apart at precise frequencies in an orthogonal manner, which enables demodulating the right frequencies. OFDM provides high spectral efficiency in the PHY layer, and is used in addition in HIPERLAN/2, xDSL (HDSL 1.6 Mb/s, ADSL 6 Mb/s, VDSL 100 Mb/s) and DAB/DVB systems.

IEEE 802.11a is a specification for a high-speed physical layer in the 5 GHz band, providing theoretical data rates up to 54 Mbps. The IEEE 802.11h standard provides 5GHz extensions in Europe where the bandwidth allocation in the 5GHz band differs from other areas in the world. IEEE 802.11g technology provides the same data rate as IEEE 802.11a, but operates in the 2.4 GHz band, thus enabling better signal penetration through walls in comparison to 5 GHz with the same transmission power. Further enhancements are due to the standards IEEE 802.11i for WLAN security, and IEEE 802.11e for providing QoS support in the WLAN environment.

The Bluetooth 1.1 specification [25] enables 721 kbps data rates, but it is also upgrading to support higher data rates. The Bluetooth Special Interest Group's (SIG) has adopted Enhanced Data Rate (EDR) specifications, which replaces the Gaussian frequency-shift keying modulation scheme used in the current Bluetooth devices with two new modulation approaches: Differential Quadrature Phase-Shift Keying (DQPSK) resulting in a 2 Mbps data rate and an eight-level Differential Phase-Shift Keying (DPSK) resulting in a 3 Mbps data rate. Further enhancements to support data rates up to 10 or 20 Mbps are due for specification version 2.0.

HIPERLAN (HIgh PErformance Radio LAN) standards [23] include four variants. Especially the HIPERLAN/2 specification provides interfaces to many kinds of networks, such as UMTS back bone, ATM and IP networks. Similar to IEEE 802.11a, HIPERLAN/2 uses the 5 GHz band and supports up to a 54 Mbps data rate. It has support for real-time data, and has special emphasis in supporting the quality of data, speech and video services. On the physical layer BPSK, QPSK, 16QAM or 64QAM modulations are used. However, HIPERLAN standard have had severe problems to enter the markets dominated by 802.11 standards. In addition, the lack of end-to-end QoS support in the public Internet so far may have lowered the interest for providing QoS support in the local access network.

As illustrated in Table 1, there are a wide variety of wireless networks, and many local and personal area standards have emerged. However, whatever the used WLAN system is, it is providing service only "locally". Vertical mobility combines the capacity of local area networks and the coverage of wide area cellular networks. In the following, integration models are briefly discussed, with focus on coupling WLAN with GPRS/UMTS.

2.2.4 Integration models

There are two models in the literature for the integration of WLAN into cellular networks, tight and loose coupling models. These two models represent the current R&D means to achieve global mobility management [26]. As illustrated in Fig. 4, the loose coupling model facilitates vertical mobility management in the IP stack, putting emphasis on implementing roaming, mobility and security management in the network layer.

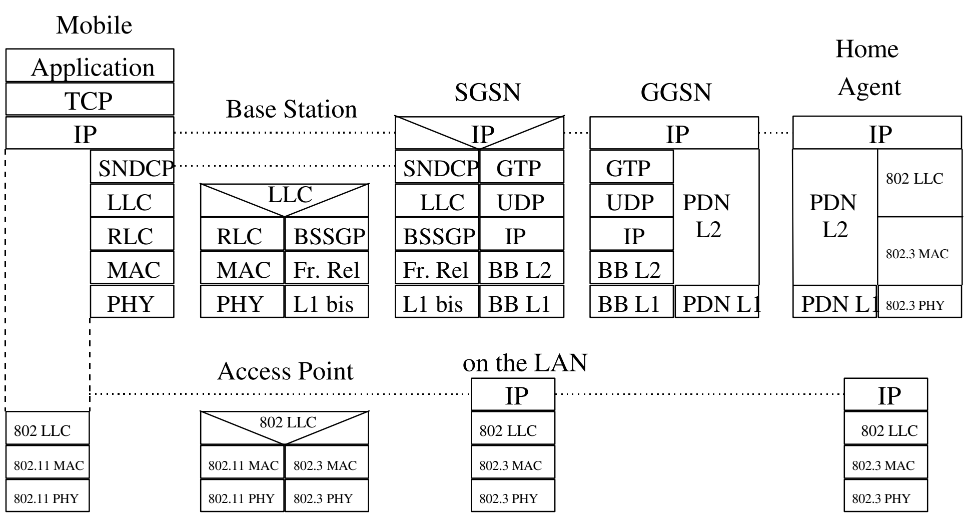


Fig. 4. Loose coupling protocol architecture of GPRS/EDGE and 802.11 WLAN.

As described in [27], vertical mobility management in the tight coupling model can be interfaced below the LLC in the GPRS signalling model, and WLAN roaming can be implemented as micro-mobility within the area of an SGSN. The Inter-Working Function (IWF) implements a thin layer of logic on top of 802.11 MAC to provide functionalities required by the LLC interface.

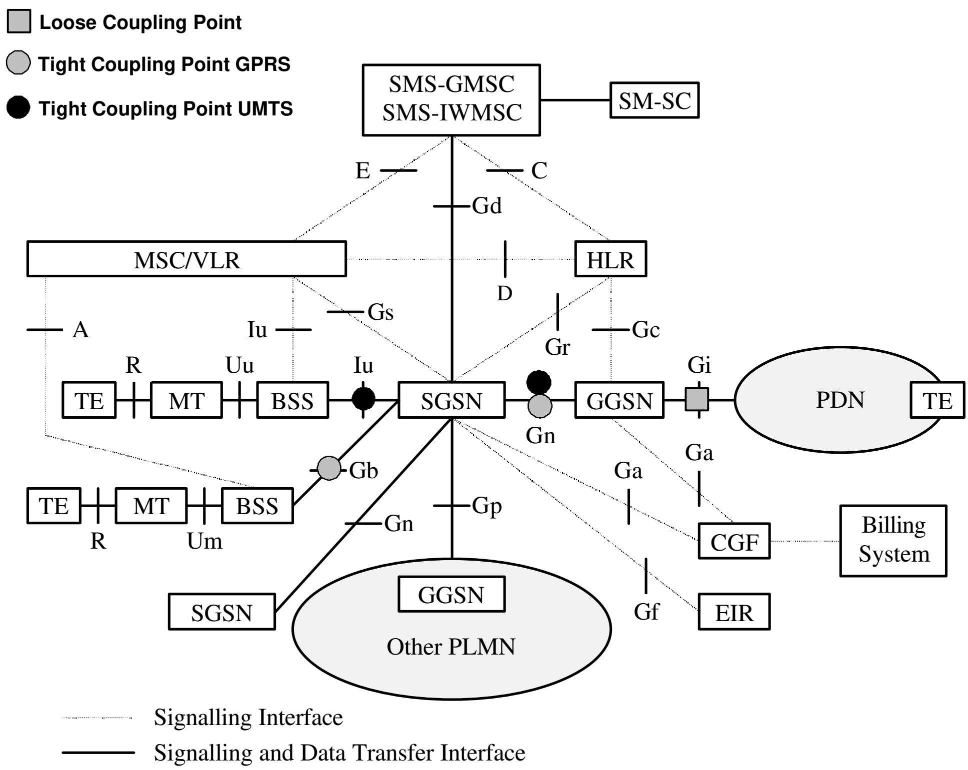


Fig. 5. WLAN loose and tight coupling points for GPRS and UMTS [29].

Fig. 5 shows an architectural view and points in the architecture for loose and tight coupling. Loose coupling uses a Gi interface, whereas GPRS tight coupling uses Gn and Gb, and UMTS tight coupling uses Gn and Iu. The tight coupling model may be more favourable for operators as the mobility, security, QoS, billing and authentication features of GPRS/UMTS are facilitated over the WLAN access channel [28].

Standardization of WLAN-Cellular interworking is undertaken by standardization groups such as 3GPP [29], 3GPP2, IETF and ITU-R. For example, the IEEE 802.11 working group has defined six scenarios for WLAN-Cellular interworking [30] including service continuity and seamless services. The GSM Association has defined 802.1X and EAP-SIM based authentication that uses a RADIUS authentication server. The reference architecture also enables an inter-operator roaming interface [31].

While vertical mobility can be considereda as primarily focused on solving roaming among cellular and wireless local area networks, there are also interworking and integrating possibilities among heterogeneous wireless networks. The first one with imminent deployment is the integration of digital audio and video broadcasts (DAB/DVB) into cellular networks [32]. Integration holds possibilities for hybrid solutions ranging from device-based integration to network-level integration [33]. In Finland there has been a pilot program to test what is called the fourth digital broadcast network, which is reserved for mobile datacasting services (IP datacast, IPDC) [34].

Another aspect for integration models is service integration. While location management functions can be seen as a part of mobility management, *location services* can be considered as an independent issue. Location services have been specified in networking related standardisation [35] and they are logically mapped to the system architecture. Location service can be internal (e.g., cell based positioning) or external (e.g., GPS based). Cell based positioning can be done in wireless local area networks as well as in cellular networks, usually involving signal measurements from several base stations. Geolocation capabilities have now been integrated into existing networks and handsets, and now one can already use location services with either cellular based services or using a GPS module and navigation software. Such has been developed, e.g., for Symbian series 60 mobile phones. The main problem with using GPS relates to battery exhaustion. Current solutions favour a modular approach through wireless Bluetooth connection. An external GPS module (i.e., “GPS mouse”) has its own power source (rechargeable battery), and the coordinate information is passed through the wireless Bluetooth data connection using standard GPS protocols.

In many academic studies it has been shown that location services can be used to aid mobility management decisions in addition to mobile navigation services. The usage of geolocation information in location aided routing schemes has been proposed for example in the context of geocasting [36] and location-aware ad hoc networking [37]. Packets can be geocasted to a specific geographic area, or an ad hoc routing protocol can use location information as a metric for routing decisions. There are also proposals for position assisted handoff algorithms [38, 39]. These examples suggest the integration of location-awareness to the integrated system architecture. In general, the combination of multi-network connectivity integrated with location technologies provides a rich area for future research studies.

The system architecture in Paper I builds mainly on top of the tight coupling model of WLAN and GPRS/UMTS, but can also be used with the loose coupling model as well. It adds GPS based location awareness to this architecture, so that network can facilitate more accurate positioning than with cell area location management. Papers Iand IV discuss about using a location-aware architecture for enabling the mobile node to be aware of WLAN hotspots in the area. When approaching, the dormant WLAN interface can be woken up when the mobile user is approaching it, and closed down when passing away from the hotspot. An example of a multi-interface device is provided in [40]. When the interface is brought up, a mechanism such as MAC layer sensing [41] can be used to help connection management.

In the next section, mobility management issues in heterogeneous systems are discussed including architectural and communication aspects, having emphasis on vertical handoff.

2.3 Mobility management in heterogeneous networks

In the following, general design principles for mobility management in heterogeneous wireless IP networks are provided by harnessing state-of-the-art mobility engineering techniques to overall design. Mobility management issues include vertical mobility scenarios, methodology, control, algorithm, procedures, protocols, metrics, parameters and performance measures. They can be roughly classified as representing aspects of methodology, methods and variables, as shown in Fig. 6.

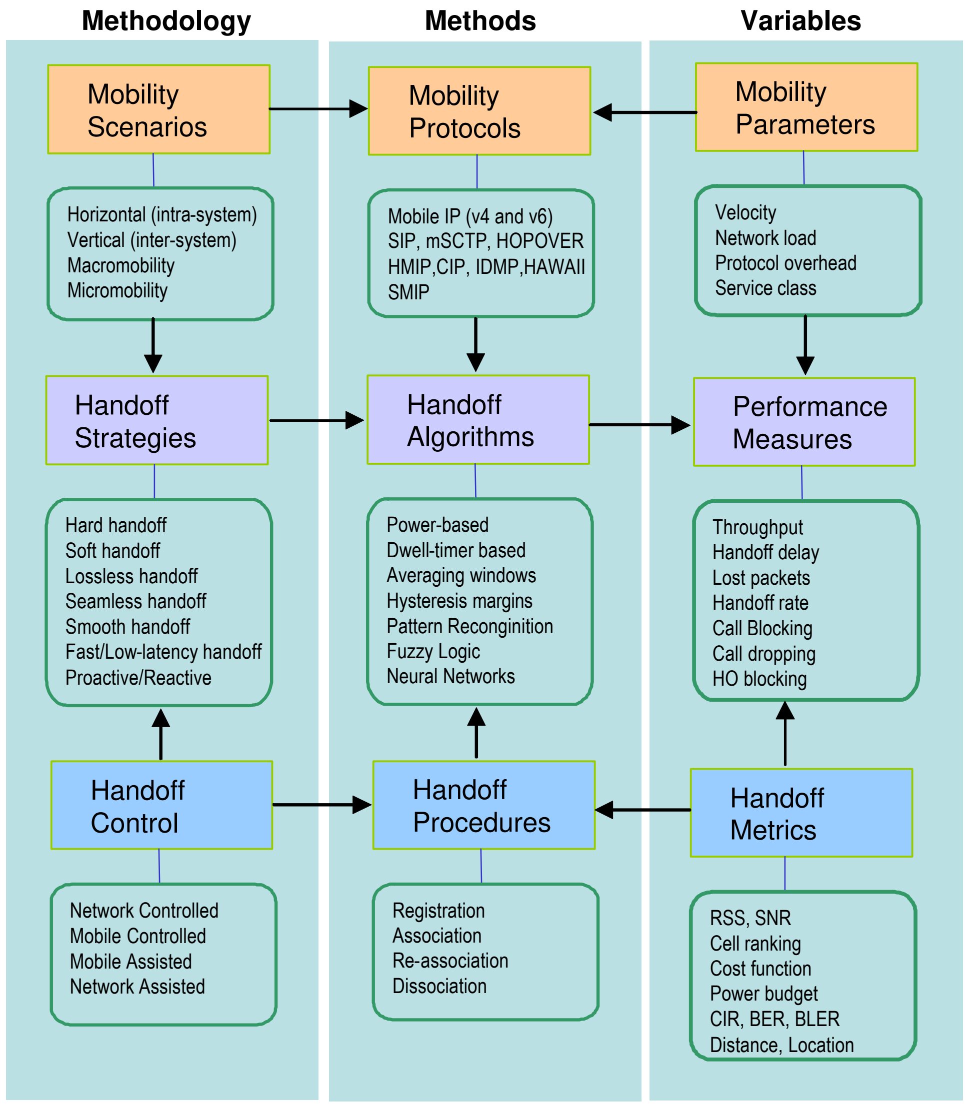


Fig. 6. Handoff and mobility management issues.

Mobility management in heterogeneous networks can take place in different layers of the OSI model including network layer (L3), link layer (L2) and cross-layer (L3 + L2) solutions [42]. In addition transport and upper OSI layer solutions exists, such as mobile middleware based on CORBA [43]. Mobility scenarios can be classified into macro- and micro-mobility, as well as into horizontal and vertical mobility. Wireless connection diversity can be interconnected using data link layer events on an L2 peer-to-peer basis [44]. System architecture design also relates to selecting appropriate mobility protocols to suite specific mobility scenarios. An overall design may utilize the best features of the available technical solutions.

2.3.1 Macro- and micromobility

Macromobility refers to mobility among administrative domains (such as an enterprise network having a number of subnetworks) and micromobility refers to mobility inside an administrative domain (such as moving from one subnet to another). Both of these require the MH to facilitate two or more network interface cards (NIC). Mobility among these NICs is also referred to as *multihoming*.

Macromobility is referred to also as *interdomain* mobility, and micromobility as *intradomain* mobility. From the integration point of view, at L3 macro-mobility can be solved with Mobile IP [45, 46] type solutions, whereas micro-mobility protocols include HMIP (Hierarchical Mobile IP) [47], IDMP (Intra-Domain Mobility Protocol) [48], Cellular IP [49, 50] and HAWAII (Handoff-Aware Wireless Access Internet Infrastructure) [49]. Fast handoff with micro-mobility becomes an issue when the MH changes is network attachment frequently, which is the case when the velocity of the MH is high or the serving area is small. Micromobility protocols also facilitate IP paging. Regional registration [51] and hierarchical schemes have been presented as partial solutions for this. A post-registration method [52] has been introduced to improve the performance of hierarchical solutions. In general, micromobility solutions enhance Mobile IP functionality by locally reducing handoff latency through minimizing the signalling load and, e.g., providing a new CoA with minimal latency. Limited multicast has been considered as a solution to improve micromobility performance [53].

Fig. 7 illustrates a vertical micromobility scenario where the BS and AP represent different RAN (Radio Access Network) technologies. In this hybrid communication scenario, multihomed mobile nodes may communicate with each other using an ad hoc protocol while preserving Internet connectivity through IP micromobility [54]. However, service stability and reachability may become an issue in such a highly mixed environment. An overall communications architecture may be needed to support different network appliances within heterogeneous RANs and administrative domains. A spontaneous communication architecture with plug-and-play features can utilize, e.g., a service gateway to connect multiple technologies and networks together [55].

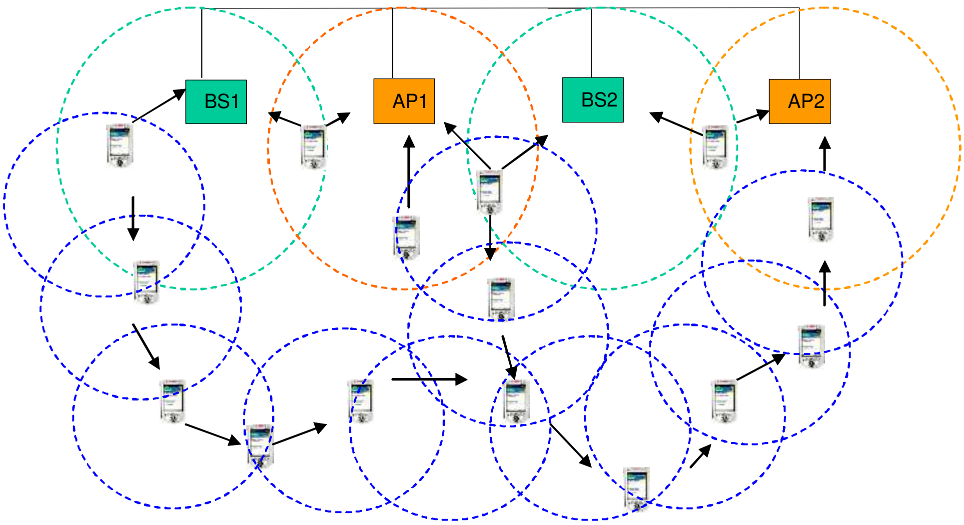


Fig. 7. Vertical micromobility with a hybrid ad hoc networking scenario.

There are a number of other proposals for macro- and micromobility protocols with certain optimized features to be considered in the overall design, especially with latency and security sensitive applications. For example, Session Initiation Protocol (SIP) supports terminal, service and personal mobility at the session layer level and can be combined [56] with Mobile IP, with possibilities of end-to-end adaptation for vertical handoff [57]. On the contrary, mSCTP (mobile Stream Control Transmission Protocol) [58] provides a transport layer solution as an alternative for L3 solutions. The Host Identification Protocol (HIP) [59] is another proposal for enabling vertical mobility above L3. It provides an additional cryptographic-based name space for Internet hosts and an API. HOPOVER (HandOff Protocol for OVERlay networks) [60] provides QoS features for Mobile IP by enabling resource reservation through wireless RSVP (ReSerVation Protocol). SMIP (Seamless Mobile IP) [61] extends Mobile IP by introducing a Decision Engine (DE) making handoff decisions for intradomain routing.

Other solutions enable *context transfer* during handoff [62]. Context information can be for example AAA parameters. Providing this type of context information prior or during the handoff enables making the new network attachment with all services in use in a seamless and secure manner. It is argued in [63] that mobile multimedia streaming in multihomed environments (i.e., the MH has a network interface to more than one administrative domain) can benefit from using *bicasting* (i.e., sending the same packet stream through different paths in order to prepare for handoff, imitating soft handoff) with Forward Error Coding (FEC) when Mobile IP handoff is used.

These examples demonstrate the versatility of technologies available to vertical roaming designers for various systems. The common denominator is that they all build on the TCP/IP model. The Internet Protocol (IP) has two versions (v4 and v6). IPv4 is currently dominant, but the transition to from IPv4 to IPv6 is expected to take place gradually [64]. It is also worth mentioning that global roaming among different 3G standards (e.g., UMTS and CDMA2000) can be achieved through IP level interworking [65].

2.3.2 Handoff strategies

Handoff (or handover) is an event when a mobile station moves from one wireless cell to another. It can be classified into horizontal (intra-system) and vertical (inter-system) cases. Horizontal handoff means handoff within the same wireless access network technology, and vertical handoff means handoff among heterogeneous wireless access network technologies. In comparison to macro- and micromobility, the terminology of horizontal and vertical reflects the wireless access network technology instead of the administrative domain (such as an IP subnetwork). Therefore there are subclasses such as vertical macromobility (mobility among different administrative domains using different wireless technologies), horizontal macromobility (mobility among different administrative domains using the same wireless technology), vertical micromobility (mobility within the same administrative domain using different wireless technologies) and horizontal micromobility (mobility within the same administrative domain using the same wireless technology).

Regardless of the mobility scenario, the decision mechanism or handoff control can be located in a network entity or in the MH itself. These cases are called *network executed* handoff (NEHO) and *mobile executed* handoff (MEHO), respectively. In GSM the measurements of alternative BSs is made by the MH, but the decision is made in the BS/MSC , thus resulting in *mobile assisted* handoff (MAHO) [66]. Another case is *network assisted handoff* (NAHO) [67] where the network can aid handoff decisions controlled by the MH.

The handoff process itself can be characterized as *hard* [68] or *soft* [69]. As it is well-know from "traditional" horizontal handoffs, hard handoff refers to a brake in the communication while making a handoff (brake-before-make), whereas soft handoff has a connection to both base stations for a while before making handoff (make-before-brake). The algorithm may be adaptive in proportion to some temporal variations in changing system parameters (e.g., changing conditions in the radio channel). These algorithms may use a cost criterion (or cost function [70]) to determine rules for optimal handoff decision.

IP mobility solutions use similar hard and soft handoff strategies, but relate the terminology to network layer phenomenon such as packet latency and packet loss. *Lossless handoff* [71] means that no packets are lost while making the handoff. *Fast handoff* [72] refers to low packet latency, which is the reason why this class is also referred to as *low-latency handoff* [73]. *Seamless handoff* [74] means that the transition to a new network attachment is transparent to the user. Application level connectivity is "kept alive". Mobile IP handoff can be *optimized* by avoiding triangle routing problem and *smoothed* through FAs using binding updates to reduce packet loss during a handoff [75]. Sometimes handoff optimization can be RAN specific [76], indicating the need to separately consider optimization of vertical, horizontal, macro and micro mobility and their variations in heterogeneous wireless access network cases.

Several aspects can be considered in the handoff decision making to optimize the handoff performance (e.g. throughput) further. One is to find rules for how and when to *trigger* vertical handoff. For this various mobility *metrics* can be employed as a specific *algorithm*. The decision about when and how this transition (or handoff) is executed is aided by *handoff policy* [77]. For example, the priority can be set to provide the fastest network connection to the mobile user, or the cheapest. Policy parameters can be weighted as a set of tuples (factor, weight) [78]. For example, WLAN can be set as high priority, so that is used always when available, or alternatively handoff is made only when no latency or communication break sensitive application is active [79]. Thus different mobility and application scenarios require different optimization aspects.

Considering the WLAN priority model, it would be preferable to seamlessly and automatically switch to a higher data rate network connection whenever it is available. This is especially important for a user that wants to download a large amount of data. Intuitively, optimization (even a sub-optimal one) of the handoff triggering instant can significantly improve the handoff performance. For a seamless handoff procedure, the delay has to be kept within certain boundaries. For a real-time user a delay of, e.g., 50 ms could be acceptable (depending on the total delay budget), whereas a non-real-time user might accept even delays as long as, e.g., 3-10 sec. Therefore the term *seamless* is also a subjective term, concluded by the user. In a situation where a terminal operates both with real-time and non-realtime applications, the delay bounds are naturally dictated by the real-time traffic. On the other hand, WLAN can be the primary network, e.g. for people who work in the office environment. In this scenario the user has a laptop or PDA in the office connected to the company WLAN. When one leaves the office, one may want to maintain some network applications running, such as a file download, connection to a database or remote terminal. At the edge of a WLAN cell, a handoff algorithm should decide when to trigger the handoff procedure and switch seamlessly to an overlaying cellular data connection.

As part of this thesis (elaborated in Chapter 3), the author has designed and implemented an experimental vertical macromobility testbed for Mobile IPv4 used with handoff algorithms such as fuzzy logic. Another test case comparison using IPv4 and IPv6 was done for a voice over IP application in WLAN system. The nature of these testbeds has been merely to demonstrate proof-of-concepts, and extensive vertical mobility test runs have not been deployed by the author, although the author has been involved with educational contribution in preparing such test cases [80, 81]. Results from the empirical and simulation studies complement the technical solutions discussed in this chapter providing more tools and knowledge for holistic system design and mobility engineering.

The author's results have demonstrated that vertical handoff delay can be very dominant even with an optimal choice of the dwell-timer value. Thus proper understanding and planning of the delay budget is necessary for seamless mobility and handoff. Different delay components (detection, configuration, registration) resulting in vertical macromobility performance with Mobile IPv6 are elaborated in [82]. It is shown that fast RA (Router Advertisement) with RA caching can benefit vertical handoff performance, especially when used with TCP level enhancements (e.g., TCP proxy). Without optimization router discovery, duplicate address detection (DAD) and the registration of new addresses are too inefficient to support seamless inter-domain handoffs in Mobile IPv6 [83]. Whether or not Mobile IP is used in vertical handoff, the handoff procedure needs elaboration. Especially, a delay budget needs to be carefully considered. Various sources of delays (packet latencies) are given in Table 4, considering both best and worst case scenarios.

Table 4. Latency budget of IPv4 and IPv6 mobility signalling with IEEE 802.11 [84].

Layer	Item	IPv4 Best Case (ms)	IPv4 Worst Case (ms)	IPv6 Best Case (ms)	IPv6 Worst Case (ms)
L2	802.11 scan (passive)	0 (cached)	1 sec (wait for Beacon)	0 (cached)	1 sec (wait for Beacon)
L2	802.11 scan (active)	20	300	20	300
L2	802.11 assoc/reassoc (no IAPP)	4	20	4	20
L2	802.11 assoc/reassoc (w/ IAPP)	20	80	20	80
L2	802.1X authentication (full)	750	1200	750	1200
L2	802.1X Fast resume	150	300	150	300
L2	Fast handoff (4-way handshake only)	10	80	10	80
L3	DHCPv4 (6to4 scenario only)	200	500	0	0
L3	IPv4 DAD	0 (DNA)	3000	0	0
L3	Initial RS/RA	0	0	5	10
L3	Wait for more RAs	0	0	0	1500
L3	IPv6 DAD	0	0	0 (Optimistic DAD)	1000
L3	MN-HA BU	0	200	0	200
L3	MN-CN BU	100	200	100	200
L4	TCP adjustment	0	Varies	0	Varies

One can note the individual sources of delays such as MAC level mobility signalling, and on the other hand IPv4 and IPv6 specific mobility signalling such as binding updates (BU) and router advertisements (RA). For optimizing the handoff decision, one should minimize the delay budget through a proactive [85] approach where as much as possible is done prior to the handoff. Then, when minimal handoff delay is enabled, the optimal handoff instant should be considered. For the decision making algorithm cross-layer information (e.g., from the PHY layer to upper layers) can be considered. The cross-layer optimization framework leading to a holistic connectivity management paradigm is described in the following.

2.3.3 Vertical handoff characteristics

The handoff procedure in general includes events such as registration, association, re-association, and dissociation [86]. The vertical handoff procedure involves these same basic events. Fig. 8 illustrates the vertical handoff procedure.

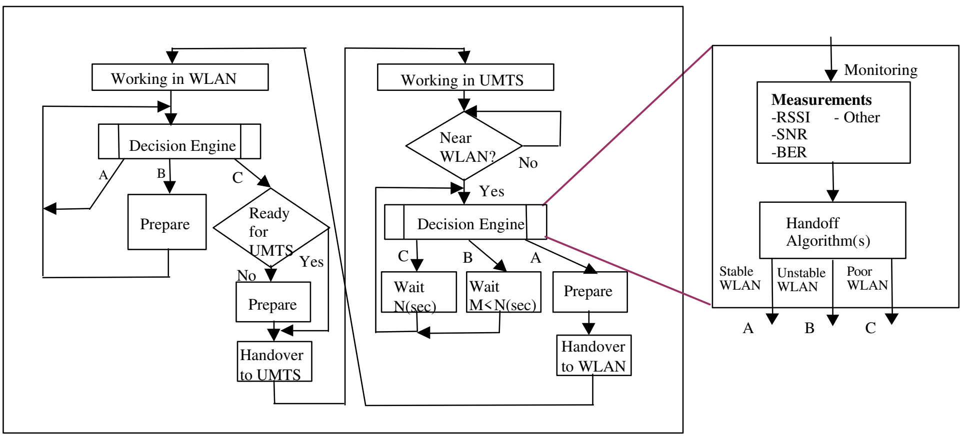


Fig. 8. Vertical handoff procedure between WLAN and UMTS.

When the MH is associated with UMTS, it monitors at repeated intervals to see whether or not a better high data rate WLAN service is available. Like suggested in Paper I, location information can aid in optimizing this interval so that it is not monitored when the MH is not close to a WLAN hotspot. In the MH there are algorithms that can be implemented as a Handoff Decision Engine (HDE) to provide rules for decision making. Fuzzy logic can be utilized to define fuzzy concepts such as “close” and “far” from the WLAN, as discussed in Paper IV. As soon as such a WLAN becomes available, the HDE should initiate an association procedure to the newly discovered AP. Depending on the HDE configuration, the MH may or may not dissociate from UMTS while roaming to a WLAN. In the tight integration model, no dissociation is taking place, only the serving point of network attachment changes. In the loose integration model using Mobile IP, corresponding relations called bindings in the binding caches are updated.

As depicted in Fig. 8, the HDE can give one of three possible action outputs: A, B and C. When the MH is associated to a WLAN, the output A (*Relax)* indicates that the link quality of the WLAN is satisfactory and there is no imminent need for handoff. Output B (*Alert)* indicates that that the WLAN signal is weakening and approaching the transition region. HDE allows the application to prepare for the coming handoff by, e.g., momentarily increasing buffer size [71] and potentially inform upper layer protocols such as to tailor the TCP congestion window size [87] or other mechanisms to improve the performance during handoff. Finally, output C (*Handoff)* indicates that a handoff is needed and the handoff procedure to the overlay network (UMTS) is invoked.

The user should also be able to configure the HDE for certain options such as “use WLAN when available”. In an ideal case the HDE would be aware of the application requirements for delay and throughput and can make decisions without user intervention.

This type of approach is suggested in Paper I where a middleware entity called Plug-and-play Application Platform (PnPAP) is utilized. The HDE would in this scenario be part of the HCon (Holistic Connectivity Management) module. PnPAP structure is illustrated in Fig. 9.

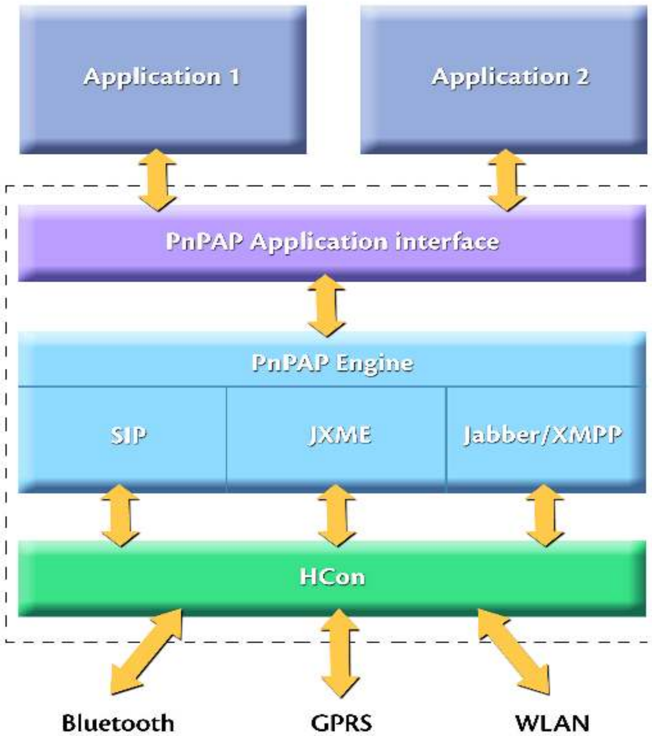


Fig. 9. HCon module in the PnPAP platform includes handoff decision engine.

The HDE monitors the selected parameters and uses them in one or more handoff algorithms. In the decision, the handoff algorithm uses the selected parameters (*metrics*) in making the decision. Handoff algorithms can then be tailored to take full advantage of the selected metrics. The selected metrics depend on the application scenario and the level where the HDE is utilized. Optionally, the HDE can use *cross-layer information* [42]. It is information that is usually not provided through standard interfaces, but crossing through OSI layers is required to form a logical channel. This can be referred to as intentional OSI layer violation, with the purpose of achieving more information to support optimal handoff decision.

In the following some of the metrics are mentioned. Handoff metrics and algorithms have been surveyed in Paper II. In the case of a traditional horizontal handoff scenario, Received Signal Strength (RSS) or its relative equivalent RSSI (Received Signal Strength Indicator) has usually been the primary monitored metric. Also the distance [88] from a BS is used in conjunction, indicating the existence of location-aware handoff algorithms. In GSM, both the MH and BS monitor the RSSI, which is a scaled number derived from

RSS, and Bit Error Rate (BER) values to know the channel quality. The availability to a middleware type of HDE implementation may be only through a relative signal quality indicator [89]. Block Error Rate (BLER) has been used to indicate MAC layer events. The probability of packet retransmission is equal to the probability of any of the transport blocks that the packets consist of [90]. The power control [91] of the MH can be another optimization criteria in conjunction with handoff, this being another metric for the algorithms to consider affecting signal quality and power consumption. In target cell determination [92] the cell selection criteria can be dependent on path loss, cell type, candidate cell capacity, or combination of them. Velocity is an important metric as it has effects on radio and network level mobility parameters [93]. In an extreme case, the MH may be travelling in a train [94]. The HDE may use in addition parameters such as network load [95] or economical factors such as price [96]. These metrics form a rich set of potential metrics to be considered in the design of HDE functionalities, enabling context-aware handoff decisions [97]. The HDE can utilize a reconfigurable state machine (as introduced in Paper I) to tailor rules for different mobility and application scenarios. The state machine can utilize or implement specific algorithms as part of the decision making.

While there is a long history of handoff algorithms in traditional horizontal handoff, not much research has been undertaken for what kind of handoff algorithm should be used in IP mobility and handoff. In many cases simplistic approaches have been used, such as a predefined threshold of WLAN RSS, where the threshold is set to a value at which the transmission and reception of IP control information is close to becoming unreliable [98]. However, article [98] points out that "the intelligence of the switching algorithm could be improved". As discussed in Paper II, based on RSS measurements a number of algorithms can be used, such as dwell-timer and hysteresis based. With the optimal value of hysteresis and averaging the handoff performance can be significantly improved [99, 100]. Practical implementations usually favour suboptimal rather than optimal algorithms due to the implementation complexity [101]. Algorithm performance can thus be compromised to some extent to favour a simpler and more light-weight implementation. Algorithm performance can be measured in proportion to throughput, handoff delay, number of handoffs and lost packets. These are mainly single-user specific measures. At the network or cell level, performance measures such as network load and blocking rates can be used [102]. This thesis studies the use of simple-to-implement algorithms based on using a dwell-timer in a vertical mobility scenario as the basis for analysis and uses mean throughput as the main performance indicator. When seeking a generic solution for handoff triggering in wireless networks, a selection or combination of some of these may be appropriate. RSS measurement algorithms are characterized by low computation complexity, but adaptive algorithms [103] have potential for more intelligent decisions improving the performance. Adaptive algorithms include pattern recognition [104], neural networks [105] and fuzzy logic [106] algorithms. The author of this thesis has designed and implemented an experimental proof-of-concept vertical macromobility testbed for Mobile IPv4 using a fuzzy logic algorithm, as presented in Paper VIII. In the following the characteristics of vertical handoff transition between cellular networks (GPRS/EDGE/UMTS) and WLAN is discussed.

This thesis considers seamless handoff for applications requiring maximum throughput, but which are not very sensitive to delay or its variations. These type of applications (e.g., downloading games or music files) could take advantage of the "passing through a hotspot" scenario, as discussed in Paper I. Passing through a hotspot involves the *transition* from cellular to WLAN ("moving-in") and vice versa from WLAN to cellular ("moving-out").

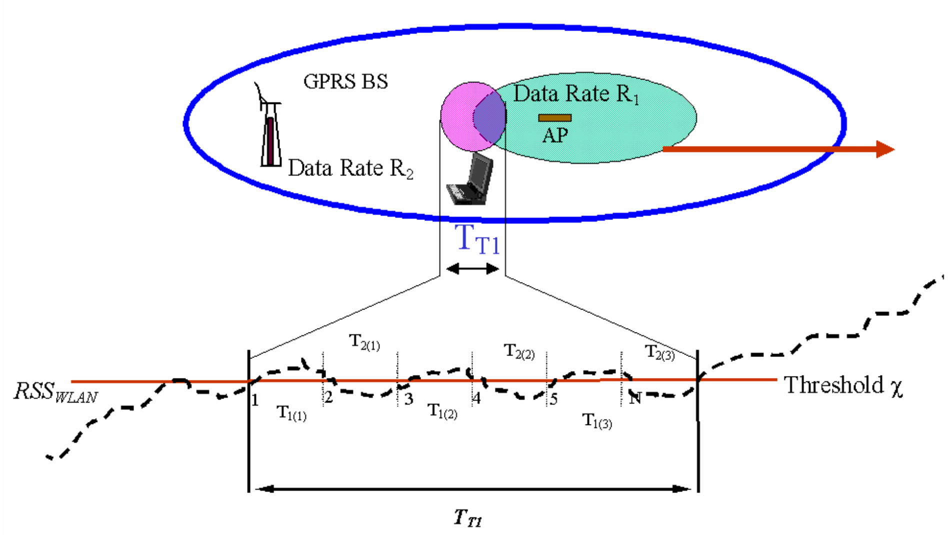


Fig. 10. Moving-in transition from cellular to WLAN.

As depicted in Fig. 10, the moving-in transition considers changing the network attachment from a GPRS BS to a WLAN AP. Depending on the handoff control scenario, a number of metrics can be involved, but generally RSS (Received Signal Strength) or its indicator (RSSI) is used in most implementations. For signal strength variations, the area that requires special attention is called the *transition region*. The transition region (T_T) refers to the time (and corresponding distance) where the measured WLAN RSS goes up and down ("dances") around the WLAN transceiver sensitivity threshold χ (such as -82 dB). In the moving-in (from GPRS/UMTS to WLAN) scenario, T_T starts from the first time instant where RSS comes above threshold χ, till the last instant where it was below the threshold. As depicted in Fig. 11, in the moving-out (from WLAN to GPRS/UMTS) scenario the transition region starts from the first time instant where received signal strength falls below threshold χ, till the last instant where it was above it.

If the two overlapping systems have significantly different effective data rates, it becomes important to utilize the system with the highest data rate (or other way highest priority system) to the maximum. For this purpose one can consider optimal and suboptimal algorithms. The latter are usually more suitable for practical implementations due to their lower implementation and computational complexity. In this thesis two simple-to-implement handoff algorithms are considered in the analysis: one based on received signal strength and one based on using a dwell-timer. These studies are discussed in Papers I, VI, VII. In addition, in the testbed demonstration a fuzzy logic algorithm was implemented, as discussed in Paper VIII. In this thesis, the performance of these RSS and dwell-timer algorithms are analysed in transition regions of both moving-in and moving-out scenarios. The performance, measured as the mean throughput (bits/s), is a function of the terminal velocity (v), the handoff delay (Δ), and the ratio of the effective data rates (Ω).

The specific interest is to analyze these two basic algorithms in the transition region in moving-in and moving-out scenarios when the network asymmetry with practical data rates is considered. This approach has limitations, but it provides a basis for more advanced studies. The transition region is prone to communication errors above PHY layers, causing lost packets and retransmission both at the data link and transport layers. However, if the ratio of the effective data rates is significant enough, that will compensate for these effects to some extent. Time spent in the transition region is not known beforehand as it is related to the direction and speed of the movement. Naturally, the time spent within the WLAN cell area before getting into the transition area is the most efficient one. The point made in this thesis is that potentially also the transition area is important to consider, especially in the case when the ratio of effective data rates is high, favouring staying as long as possible in the higher data rate system. This thesis has studied how long of a dwell-timer is profitable in such cases when assuming different amounts of handoff delay, the ratio of effective data rates and the reduction to nominal rates for a single user.

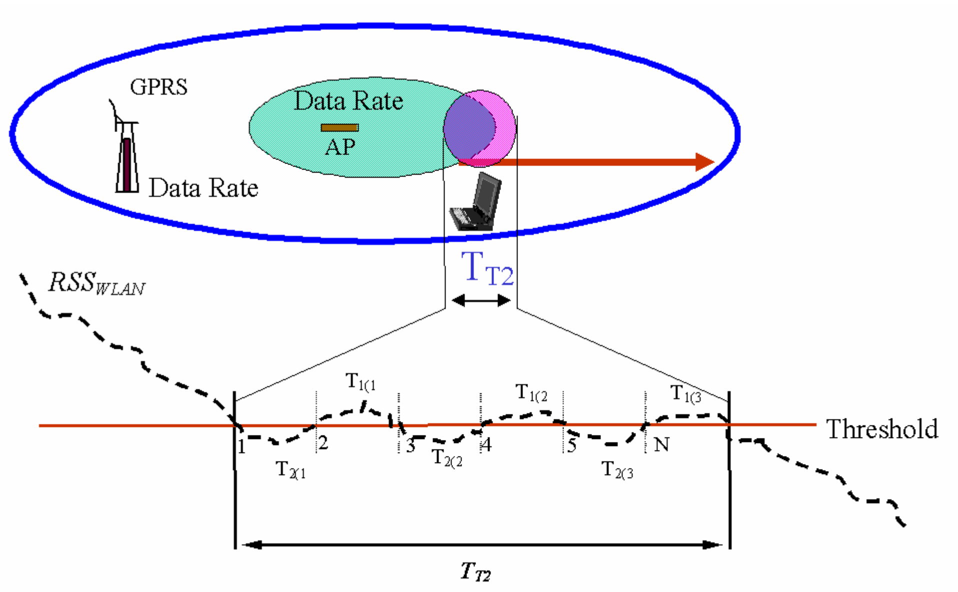


Fig. 11. Moving-out transition from WLAN to cellular.

While there are location-aided handoff algorithms [107], location-awareness is considered in this thesis only in the architecture when preparing the MH for approaching vertical handoff. Due to the fact that the RSS variation at the edge of the WLAN cell is very fast, GPS based location information assumed in the architecture may be too sensitive to delay and estimation errors to be used in the handoff decision itself. In the transition region signal fading is fast, and location information would need to be very accurate and low-latency in order to facilitate performance improvement. From this aspect, practical usage scenarios for location-aware vertical mobility management may exist merely in macro-environments rather than in micro-environments.

A holistic understanding of what happens in the transition region should be helpful for any programmer or system designer considering vertically overlapping access networks and seamless mobility across them. This is one of the main purposes for the simulations and transition analysis presented in this thesis. Vertical mobility also has relations to resource allocation and quality-of-service aspects, which are discussed in the following.

2.3.4 Resource allocation aspects

There can be capacity sharing among heterogeneous wireless networks, especially when the tight coupling model is considered. For the operator, vertical roaming provides, in some network set-ups, means for network off-loading from GPRS to WLAN [27], thus making vertical handoff a resource allocation instrument. Resource allocation in cellular networks includes tasks such as communication Channel Admission Control (CAC), channel assignment, power control and handoff, adapting to the time and space variant changes in signal quality, network load and other environmental variables [108].

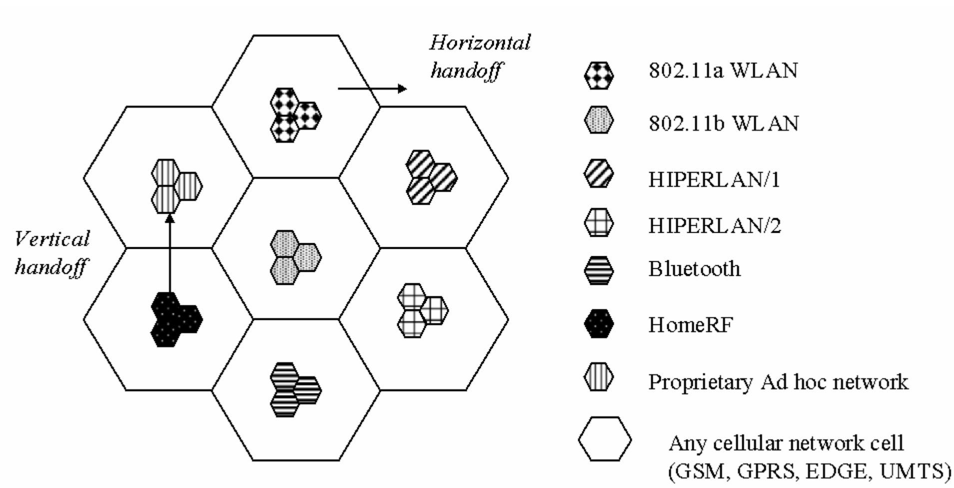


Fig. 12. Vertical mobility model.

The vertical mobility model in Fig. 12 illustrates how several independent micro-cells (representing different wireless access technologies) can be overlaid by macro-cells. Both vertical and horizontal handoffs can take place in this mixed and heterogeneous overlay system. Vertical handoff related resource allocation aspects are due to the nature of vertically overlapped asymmetric networks. They facilitate different features for available effective data rate (or throughput) and delay characteristics. They also differ in the way physical and data link layer communication takes place. Cellular networks facilitate resource allocation by the means of logical channel assignments, whereas WLAN systems (such as IEEE 802.11b) rely on collision avoidance mechanisms at the MAC layer. In order to facilitate resource allocation over hybrid networks, resource allocation and logical channel allocation is required throughout the vertically integrated system.

From the modelling point of view, handoff processes (whether they take place in the horizontal or vertical plane) can be modelled with multidimensional birth-death processes [109]. Changes in the handoff process can be seen as variations in system states. Following the formulation logic and principles presented in [110], state v_{gz} of a single cell in the vertical mobility model is a sequence of integers, which can be written as Z n-tuples forming a multi-dimensional state space of dimension Z·W:

$$
\begin{array}{cccccc}
v_{11}, & v_{12}, & \cdots & v_{1w}, & \cdots & v_{1W} \\
v_{21}, & v_{22}, & \cdots & v_{2w}, & \cdots & v_{2W} \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
v_{z1}, & v_{z2}, & \cdots & v_{zw}, & \cdots & v_{zW} \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
v_{Z1}, & v_{Z2}, & \cdots & v_{Zw} & \cdots & v_{ZW}
\end{array}
$$

State space variables include Z platform types, indexed by z = 1, 2, 3, …, Z (e.g. classified by mobility type or speed: stationary, pedestrian, vehicle, …), W call or traffic types, indexed by w = 1,2,3, …, W (e.g. voice, data etc.), R_k resource types to support calls, k = 0,1,2, …, K-1. Platform z can support one of the W call types at any time, and a single platform may also use multiple channels. States v = 0, 1, 2, …, v_{max} and corresponding state probabilities p(v) can be calculated. To find statistical equilibrium cell state probabilities, flow balance equations need to be defined. The procedure for calculating theoretical resource allocation performance includes identifying first permissible cell states, relating the parameters of the underlying driving processes to the state probability transition rates, determining the state probability flow balance equations and solving the resulting set of simultaneous nonlinear equations for the state probabilities. From these the performance measures such as handoff blocking probability and forced termination probability can be calculated. *Dwell-time* refers to the duration of time MH spends in one cell, whereas the *dwell-timer* is a handoff algorithm parameter to adjust how long the MH will persist in a specific cell. Dwell-time can be modelled as a system level parameter [111]. Application session duration is another parameter to consider, as discussed in [110]. It is commonly modelled using a negative exponential distribution with parameter μ. This distribution may not be the best for data traffic. For example, SOHYP (Sum of Hyperexponentials) distribution can be used instead. It can be used to approximate the behaviour of any positive random variable.

The basic model for a handoff arrival process applies to connection-oriented systems, including those which use circuit switching, virtual circuit switching (e.g. TCP or RTP/VoIP applications) as well as systems which employ admission control (e.g. HIPERLAN/2 or any WLAN with a CAC mechanism). If CAC is to be used together with random access WLAN systems, a multiple radio network environment can defined having a total of C logical traffic channels consisting of, e.g., C_R channels allocated to real-time traffic (e.g.,voice) and C_N channels to non-realtime (e.g., WWW) traffic. In addition there can be allocated C_C control channels and C_H handoff cut-off priority channels [110]. Performance measures in this type of analysis include blocking probability (i.e., fraction of new call attempts that are denied access to a channel in their cell) and handoff failure probability (i.e., fraction of handoff attempts that fail to get a channel in their target cell). It is assumed that some number of channels C_H are reserved for arriving handoff calls (cut-off priority).

Additional performance measures include forced termination probability (i.e., probability that call a is interrupted due to a handoff failure during its lifetime), handoff activity factor (i.e., expected number of handoff attempts for nonblocked calls), carried traffic per cell (i.e., average number of channels occupied by the calls from the given platform type) and statistical measures such as equilibrium state probabilities and average handoff call arrival rate. This is a more theoretical aspect to vertical mobility from the aspect of resource allocation, and requires a more theoretical and mathematical approach than was applied in this thesis. However, this is an important area for more complete system level studies. For example, resource allocation provides the means to handle differently real-time and non-real-time application demands during handoff. A handoff prioritization scheme can utilize, e.g., pre-emptive priority handoff [112] to balance between real-time and non-realitime application blocking and forced termination probabilities. The resource allocation problem in heterogeneous wireless networks is an area for much future study, and new paradigms such as considering game theoretic approaches [113] and information theory [114] can be considered.

For next generation wireless networks, one of the key concepts is the Quality of Service (QoS) that is currently provided mainly on a *best effort* basis. Especially delay sensitive applications require end-to-end quality of service management [115]. Selections among heterogeneous wireless networks should consider also the QoS that the given wireless network can provide to enable the *always best connected* [116] connectivity management paradigm instead of the current best effort. For resource allocation in IP networks, there are two approaches: differentiated services (Diff-serv) and integrated services (Int-serv). Diff-serv [117] enables service differentiation by marking the IP packets in the Diff-serv field (in both IPv4 and IPv6), thus enabling routers to differentiate among different types of traffic (real-time, non-realtime). On the other hand, the Int-serv approach may employ a specific resource reservation protocol such as RSVP to directly steer the resource reservation process [118]. There are also proposals for co-operation [119] of these two approaches, especially in tailoring the performance in the case of frequent handoffs. Enabling end-to-end QoS means that all used RANs support quality of service and traffic differentiation also at the MAC layer. For the de-facto WLAN standard IEEE 802.11 this has been only recently introduced in the form of the new IEEE 802.11e [120, 121]. One has to also consider how different packet sizes and class-based quality of service mechanisms survive over the wireless radio air interface [122]. For avoiding protocol overhead in the wireless transmission, header compression [123] must be used over the last hop communications. In addition, the codecs used at the application level can adapt to the varying conditions at the lower layers of communications [124]. Holistic system design should consider all of these aspects in order to pave the way for developing advanced mobile wireless applications, making the application development process faster [125]. With clear API definitions to intelligent middleware, application and service developers can utilize the underlying resources in an efficient and seamless manner, resulting in better and more numerous mobile applications in a long run.

This chapter has provided an overview of the state-of-the-art technologies with pointers for further study. It has discussed the enabling technologies for vertical mobility management and discussed how this thesis work contributes to this field. It has been seen that understanding and optimizing the handoff process is an integral part of the holistic system design. One of the key challenges in future network management is end-to-end optimisation that takes into account wireless optimizations, routing optimizations and delay budget for heterogeneous wireless environments. An important area for traffic optimization in these next generation wireless networks is to manage the mean throughput and the delay in the handoff process between heterogeneous networks. The next chapter summarizes the contributions in the original papers of this thesis.

3 Summary of contributions in the original papers

"Science is not a mechanism but a human progress; and not a set of findings but the search for them." - J. Bronowski

In this chapter a summary of contributions in the original papers is given. As indicated earlier, the thesis is related to the analysis of the mobility and handoff in heterogeneous networks at architectural, protocol and algorithm levels. Step by step, looking at the research problem from different angles, the thesis establishes an analytical framework and holistic architectural design for implementing vertical mobility. Contributions in the individual papers are incremental pushes towards making human progress in this research area.

3.1 System architecture and mobility management

The first four articles in this thesis contribute to the system design aspects of vertical mobility. Paper I, entitled "Analysis of Handoff in Location-Aware Vertical Multi-Access Networks", presents an architecture for seamless location-aware integration of WLAN hotspots into cellular networks. The author had the main responsibility for the technical content and simulation results. The integration of WLAN hotspots into next generation cellular networks requires considerations on location management, resource allocation, handoff algorithms and their sensitivity to mobility related features such as velocity of the mobile and the handoff delay. For a mobile station, it may be beneficial to know the whereabouts of the hotspots in order to facilitate optimal handoff between the two access networks.

In an integrated cellular-WLAN environment, one has to consider location management as part of mobility management due to differences in the cell sizes of the two overlapping technologies. Location information acquired through GPS may become helpful in some cases for advanced location management. In this paper a location-aware architecture is proposed to support vertical roaming among heterogeneous wireless access networks. The article includes a description of the system architecture and the procedures and algorithms needed to implement mobility and location management. An analysis is provided for locally optimal handoff decision when moving in and out of a hotspot. A comparison is given of two handoff algorithms (power and dwell-timer based) in moving-in and moving-out transitions, and their sensitivity to mobile velocity and handoff delay.

Paper II, entitled "Handoff in Hybrid Mobile Data Networks", is an example of early work in this area. The author of this thesis was a co-author in this paper, contributing to the technical content in the areas of analyzing GPRS based mobility management and related comparisons. The paper presents an overview of issues related to handoff, including a survey of algorithms and architectural issues. Five architectures for the example hybrid network, based on the emulation of GPRS entities within the WLAN, Mobile IP, a virtual access point, and a mobility gateway (proxy) were described and compared. This paper has been widely cited.

Paper III, entitled "Handoff Procedure for Heterogeneous Wireless Networks", presented a network layer Mobile IP based handoff procedure between WLAN and GPRS as a case study example. This was one of the first papers to elaborate the handoff procedure itself, and the author of this thesis was the main author in this paper. The paper emphasized added flexibility of heterogeneous wireless networks in the form of robust inter-technology mobility management schemes and sophisticated algorithms. This involves the identification of procedures, algorithms and metrics involved in handoff in heterogeneous wireless networks, and the role of the IP protocol when designing the system architecture.

Paper IV, entitled "Geolocation Information and Inter-technology Handoff", was one of the first papers to elaborate using location information in vertical handoff (referred to then as Inter-technology handoff). The paper also identified the importance of location-awareness as a trend towards next generation telecommunication systems, having potentially an effect on network management design issues. As the cellular network can be overlaid with a scattered high data-rate WLAN network, the terminal needs some information about when to search for WLAN coverage. Considering a scenario where the search is done automatically instead of manually, to avoid the unnecessary search for WLAN beacons the terminal should be aware of the whereabouts of the overlay system to be visited. This paper, where the author of this thesis was the main author of the technical content, suggested the usage of geolocation information in mobility management via distributed location databases to enable a moving host to prepare for handoff. Three scenarios and the related problems were presented, and the requirements for the system architecture and terminal were outlined. Considerations on the algorithm were given and preliminary simulation results for a fuzzy logic handoff algorithm were presented.

While these four papers had an emphasis on system design, they also contained simulation results. Especially Paper I concluded the simulation studies using the refined analytical framework that is discussed in the following.

3.2 Simulation results

The objective of the simulation results was to establish a multi-dimensional variable space. As illustrated in Fig. 13, the variables included mean throughput, velocity, ratio of effective data rates, handoff delay and dwell-timer value. Paper V, entitled "Analytical

Framework for Handoff in Non-Homogeneous Mobile Data Networks", provided a basis for the simulation model. The paper paid attention to the handoff decision among networks having different nominal data rates. The asymmetry in the data rates is the fundamental difference that calls for particular handoff strategies for non-homogeneous mobile data networks. Preliminary results were based on a combination of analysis and simulation. The author of the thesis contributed to the technical content of this paper, and continued to work with the simulation model to get more quantitative results using additional parameters in the model.

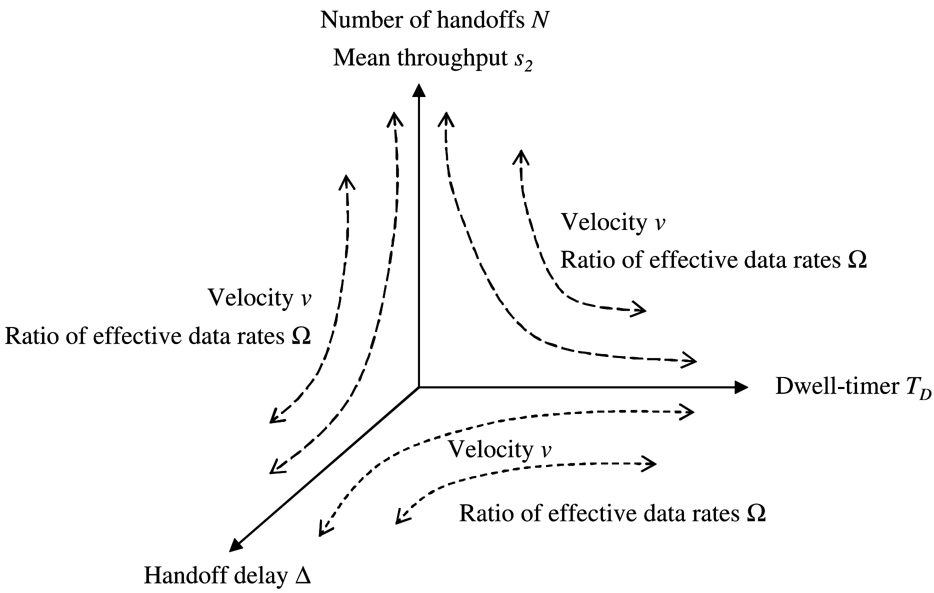


Fig. 13. Simulation results multi-dimensional variable space.

The first results of this continual work were published in Paper VI entitled as "Optimization Scheme for Mobile Users Performing Vertical Handoffs between IEEE 802.11 and GPRS/EDGE Networks". This paper presented simulation results for mean throughput and handoff delay obtained in vertical handoff and horizontal handoff in IEEE 802.11 and GPRS/EDGE networks. An optimization scheme based on finding locally optimal dwell-timer value for mobile users performing vertical handoffs was presented with analysis. The author of this thesis was the main author of this paper with core contribution to the technical content and simulation results.

Paper VII, entitled "Supporting Resource Allocation with Vertical Handoffs in Multiple Radio Network Environment", continued further developing the analytical framework towards multiple (more than 2) overlapping systems. The author of this thesis was the main contributing author to the technical content and simulation results. One of the main contributions in this paper was the formulation and generalization of handoff performance between multiple radio networks with different data rates which is given with analysis. This paper also identified the vertical handoff problem as a resource allocation support method. Resource allocation in this kind of heterogeneous environment has not been yet fully addressed in the current research work. Clearly, additional bandwidth offered by wide-band local area wireless networks opens a new area for network performance optimization. For that, it is necessary to understand the processes that have an effect to the system performance. In particular, vertical mobility is usually characterized by handoffs between systems having different data rates. The most critical area for system performance is at the very edge of the cell, where the RSS varies around the sensitivity threshold of the receiver. Simulation results showed estimates of the delay and general performance in a multiple radio network environment.

3.3 Empirical results

Paper VIII is entitled "Inter-technology Mobility Testbed". This paper gives an example for testing inter-technology mobility in a simple testbed. To make the switching between different networks and the connection management as flexible as possible, novel approaches must be considered. Results show that IP layer mobility management provides an efficient way to interconnect heterogeneous packet oriented networks. Simple hysteresis based and more complex fuzzy logic based algorithms were implemented into the handoff daemon program. It was seen that the fuzzy logic based algorithm performs better in the form of reduced handoff rate and optimized WLAN range usage. The author of this thesis implemented the testbed in his diploma thesis and continued that with applying fuzzy logic algorithm to the handoff daemon program to trigger handoff.

Paper IX is entitled "Comparative Analysis of VoIPv4 and VoIPv6 in a Bandwidth-limited Wireless LAN Testbed". In this paper it is analyzed how wireless VoIPv6 works in a bandwidth-limited wireless environment that does not support header compression, in comparison to VoIPv4. This was studied in a bandwidth-limited WLAN testbed. Comparisons were done with both varying ICMP payload and with actual VoIP applications with three codecs. It was seen that the delay and the jitter are more dominated by the choice of the codec (packet size) than the choice of IP protocol version. The differences with IPv4 and IPv6 are not big when it comes to the effect of header overhead over a last hop. There is a tradeoff between the codec (voice quality, data rate, compression algorithm complexity) and network utilization, but the overhead of VoIPv6 is not a problem in comparison to VoIPv4. The author of this thesis took part and managed the implementation work as part of his licentiate thesis, and analyzed the results.

Paper I is entitled "Plug-and-Play Application Platform: Towards Mobile Peer-to-Peer" and is the most recent publication, based on Application Supernetworking (All-IP) project themes. The author of this thesis has been instrumental in designing intelligent mobile middleware called the Plug-and-Play Application Platform (PnPAP). Specific interfaces have been defined with the project group. While Peer-to-Peer (P2P) has emerged as a new hot communication concept among Internet users, mobile usage of P2P applications is still taking its first steps. This article first elaborates the evolutionary process that P2P architectures are going through. Challenges and requirements for mobile P2P are identified, followed by a definition of a novel Plug-and-Play Application Platform (PnPAP). This platform enables dynamic selections between diverse P2P and session management protocols while preserving the best available network connectivity through Holistic Connectivity (HCon) management. On-the-fly reconfiguration and run-time parameter optimization can be done with a lightweight interpretable state machine. The concept enables flexible and seamless communications for mobile devices in P2P networks.

These original papers are reprinted at the end of the printed version of this thesis, with the permission from the original publishers. In the electronic version they are not included, but are available via the original publishers. In the next chapter, conclusions are drawn from the overall work.

4 Conclusions and future work

“The man of action has to believe, the inquirer has to doubt; the scientific investigator is both.” - C.S. Pierce

Vertical mobility has been a topic of research for over a decade now, and it is deployed in commercial products and field tests. Yet, the popularity of seamless services has not taken its place in the every day life of consumers in the same way that talking to a mobile phone or using Internet from a home PC. While there is some doubt if vertical mobility will ever have significant enough revenue creating ability for operators, the future challenge is to “put into action” services and applications that utilize vertical roaming with both technical and economical excellence. These services need to be enabled and introduced in mobile handsets with viable and tailored applications in order to see their full benefits. The challenge is about enabling better mobile applications through holistic plug-and-play connectivity and “always best connected” paradigms.

This thesis presents a holistic approach for system architecture design for the seamless location-aware integration of WLAN hotspots into cellular networks. Furthermore, it provides a transition analysis of making a seamless and optimized vertical handoff in moving-in and moving-out scenarios, taking full benefit of the asymmetric data rates provided by the heterogeneous wireless network. The thesis demonstrates various system and mobility engineering design options including optimal dwell-timer settings, protocol suite selections and location-aware architectural design. The proposed location-aware architecture enables the mobile node to prepare for approaching vertical handoffs and to wake-up the hotspot interface. The needed communication procedures between network nodes were discussed, and inter-related issues of mobility and geolocation information were considered in proportion to usability, advantages and limitations.

It is seen that holistic vertical system architecture design requires considerations on not only providing the logical infrastructure but also optimized functionality through a cross-layer paradigm in handoff decisions. While mobile applications develop towards peer-to-peer capabilities, it is needed to implement handoff decision engines and corresponding algorithms to mobile middleware in a cost-efficient way enabling light-weight implementation. Application programmers and system designers get a set of heuristics for tailoring their design and also let mobile users configure applications to meet their needs in various usage scenarios including macro- and micromobility.

The thesis introduces a framework for the analysis of vertical handoff algorithm sensitivity to various mobility parameters including velocity, handoff delay and dwell time. The usage of a variable length dwell-timer was analyzed as one potential scheme for optimizing vertical handoff locally in proportion to mean throughput perceived by a single user. While the dwell-timer is well-know from its use for handoff in homogeneous networks, it was needed to relate its usage in asymmetric non-homogeneous networks to key performance metrics and to find locally optimal values. Simulation results showed that the usage of a dwell-timer was justified when the ratio of the effective data rates was significant enough, e.g., in a situation when either network is experiencing congestion due to high network load. The dwell-timer algorithm was compared to a power based algorithm to find out how sensitive they are to the changes in effective data rates, velocity of the terminal and the amount of handoff delay. It was seen that the dwell-timer algorithm was less sensitive to the increase of velocity in comparison to the power based algorithm. This was especially the case in the moving-out scenario.

The moving-in and moving-out scenarios do not seem to be equivalent due to the post transition effect in the moving-out scenario. The dwell-timer algorithm seemed to be less sensitive to the increases of handoff delay in the moving-out scenario than in the moving-in scenario. The optimal value of the dwell-timer was varying depending on the scenario and selected simulation parameters for the MH velocity and handoff delay, thus indicating a need for an adaptive dwell-timer to adapt to these varying conditions. At least some sort of run-time configuration could be considered. This could be achieved through state-machine implementation that facilitates run-time reconfiguration. The optimal dwell-timer parameters could be communicated with the presented architecture (e.g., from BSC to MH prior to the handoff based on statistical observations). The optimal dwell-timer value varies from a minimum sampling value (100 ms) to several seconds. These two core algorithms can be integrated with more handoff decision logic for joining and leaving hotspots. For a real application one has to consider the upper layer performance as well. For example, the TCP congestion window size should be adjusted. Clearly, the usage of a dwell-timer requires upper layer protocol adaptation, leading to cross-layer approach.

The analytical framework was also extended to multiple radio network environments. Handoff profitableness was formulated and rules for decision making or analysis of the profitableness of the decision that has been made were given. It was shown that handoff delay is dominant even with an optimal choice for the dwell-timer. For optimizing the handoff decision, one should minimize the delay budget through a proactive approach where as much as possible is done a priori the handoff. Then, when minimal handoff delay is enabled, the optimal handoff instant should be considered.

Experimental results were related to a prototype system for the test, measurement, and analysis of the behaviour of the protocols used in wireless IP networks. For vertical mobility management WLAN is the primary overlay technology. It is important to understand the operational environment and radio and performance characteristics at the edge of the WLAN cell. Vertical handoff was not considered in this case, but the experimentation with a wireless voice over IP application helped to understand the performance and bandwidth requirements of voice over IP application in proportion to selecting between IPv4 and IPv6. The benefits of IPv6 will eventually boost the utilization of VoIP, as all-IP paradigms suggests that there will be a large number of wireless IP devices that need global addresses. The need for header compression, however, becomes more evident when lower bit rate codecs are used, as the proportional protocol overhead increases.

Prototype systems demonstrate the results of using Mobile IP with a fuzzy logic algorithm for vertical handoff in a heterogeneous network environment. The proof-of-concept implementation demonstrated that IP layer vertical macromobility management provides an efficient way to interconnect heterogeneous packet oriented networks. Fuzzy logic algorithm showed potential to be well suited for multiple-metric (multi-criteria) handoff decision. Classical algorithms could achieve the same results using several thresholds and processing the inputs with a series of if-then directives, but are more complex to design.

Latest contributions include developing plug-and-play middleware functionalities for Symbian mobile devices, extending the use of the earlier results to state-of-the-art mobile devices. The implementation of vertical mobility in Symbian and other mobile handsets is an interesting area of future development. This leads to using mobile middleware solutions. The Plug-and-Play Application Platform (PnPAP) enables seamlessly using different communication protocols and connectivities. Holistic Connectivity (HCon) management is part of intelligent middleware that uses lightweight state machines for run-time parameterisation to provide rules for switching connectivities and to trigger a vertical handoff process when needed. The state machine can be designed with a graphical tool where various decision rules discussed in this thesis can be implemented.

Future work includes further considerations of holistic connectivity management that takes into consideration further aspects of the cross-layer approach, rules for mobility and session management, combined radio and network resource management and adaptive QoS control. Hybrid networking requires a top-down approach that takes into account novel communication concepts such as application supernetworking with new lead application scenarios in the highler layers of the OSI model, and hand-in-hand provisioning tailored wireless and wired network services on the lower OSI layers to meet the application and service requirements. For these aspects the technical solutions presented in this thesis open the door to the future only partially. They can be considered as a beginning of more extensive studies, not the final answer. The stage is open for new technological innovations: by learning from the lessons of the past; by taking advantage of existing knowledge and knowhow; and ultimately, creating something genuine.

References

1. Stemm M & Katz R (1998) Vertical handoffs in wireless overlay networks. Mobile Networks and Applications 4: 335-350.
2. Perkins CE (1992) Simplified routing for mobile computers using TCP/IP. Proc. 1st IEEE Conference on Wireless LAN Implementation, Dayton, Ohio, USA, 1: 7-13.
3. Myles A, Johnson DB & Perkins C (1995) A mobile host protocol supporting route optimization and authentication. IEEE J Sel Areas Communic 5: 839-849.
4. Baker M, Zhao X, Cheshire S & Stone J (1996) Supporting mobility in MosquitoNet. Proc. USENIX Technical Conference, San Diego, California, USA. Cited October 19th 2004 from http://www.usenix.org/publications/library/proceedings/sd96/baker.html
5. Evans BG & Baughan K (2000) Visions of 4G. J Electronics & Communication Engineering 6: 293-303.
6. Suk YH & Kai HY (2003) Challenges in the migration to 4G mobile systems. IEEE Communications 12: 54-59.
7. Raivio Y (2001) 4G - hype or reality mobile communications. Proc. 2nd IEE International Conference on 3G Mobile Communication Technologies, London, UK, 1: 346-350.
8. Larson LE (2003) The impact of emerging 4G systems on the performance and complexity requirements of RFICs. Proc. IEEE Radio Frequency Integrated Circuits Symposium, Philadelphia, Pennsylvania, USA, 1: 15-18.
9. Craninckx J & Donnay S (2003) 4G terminals: how are we going to design them? Proc. 40th IEEE Design Automation Conference, Anaheim, California, USA, 1: 79-84.
10. Pereira JM (2001) Balancing public and private in fourth generation. Proc. 12th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, San Diego, California, USA, 2: 125-132.
11. Bargh MS, Laarhuis JH & Plas D-J (2003) A structured framework for federation between 4G-service platforms. Proc. 14th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Beijing, China, 3: 2626-2630.
12. Varshney U & Jain R (2001) Issues in emerging 4G wireless networks. Computer 6: 94-96.
13. Safwat A (2003) A-cell: a novel multi-hop architecture for 4G and 4G+ wireless networks. Proc. IEEE Vehicular Technology Conference Fall, Orlando, Florida, USA, 5: 2931-2935.
14. Xiaoxin Wu, Mukherjee B & Ghosal D (2004) Hierarchical architectures in the third-generation-cellular Network. IEEE Wireless Communications 3: 62-71.
15. Brewer EA, Katz RH, Chawathe Y & et al (1998) A network architecture for heterogeneous mobile computing. IEEE Personal Communications 5: 8-24.
16. Salkintzis AK (2004) Interworking techniques and architectures for WLAN/3G integration toward 4G mobile data networks. IEEE Wireless Communications 3: 50-61.
17. Mähönen P, Riihijarvi J, Petrova M & Shelby Z (2004) Hop-by-hop toward future mobile broadband IP. IEEE Communications 3: 138-146.

18. Ibnkahla M, Rahman QM, Sulyman AI, Al-Asady HA, Jun Yuan & Safwat A (2004) High-speed satellite mobile communications: technologies and challenges. Proc. IEEE 2: 312-339.
19. Zhang C, Xu X & Lin X (2002) Stratospheric telecommunication in 4G Communications. Proc. IEEE International Conference on Circuits and Systems, Scottsdale, Arizona, USA, 1: 172-176.
20. Hämäläinen J (1996) Design of GSM high speed data services. Ph.D thesis, Tampere University of Technology Publications 186.
21. Brasche G & Walke B (1997) Concepts services and protocols of the new GSM Phase 2+ General Packet Radio Service. IEEE Communications 8: 94-104.
22. Korhonen J, Aalto O, Gurtov A & Laamanen H (2001) Measured performance of GSM HSCSD and GPRS. Proc. IEEE International Conference on Communications, Helsinki, Finland, 5: 1330-1334.
23. Yeh J, Chen J & Lee C (2003) WLAN standards. IEEE Potentials 4: 16-22.
24. Mikkonen J (1999) Quality of service in radio access networks. PhD thesis, Tampere University of Technology Publications 255.
25. Petrioli C, Basagni S & Chlamtac M (2003) Configuring BlueStars: multihop scatternet formation for Bluetooth networks. IEEE Trans Comp 6: 779-790.
26. Salkintzis AK, Fors C & Pazhyannur R (2002) WLAN-GPRS integration for next-generation mobile data networks. IEEE Wireless Communications 5: 112-124.
27. Ekberg J & Yang F (2002) GPRS over WLAN access. Mobile Internet Technical Architecture 3: 289-296.
28. Shi M, Shen X & Mark JW (2004) IEEE 802.11 roaming and authentication in wireless LAN/cellular mobile networks. IEEE Wireless Communications 4: 66-75.
29. 3GPP (2003) Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking. 3GPP TR 22934 (Release 6) Technical Specification V620.
30. Pazhyannur RS (2003) WLAN-Cellular Interworking. IEEE 80211-03/193r1, Cited October 19th 2004 from http://wwwieee802org/11/Documents/
31. GSM Association (2004) WLAN Roaming Guidelines. Official Document IR61 v 310.
32. Benali O, El-Khazen K, Garrec D, Guiraudou M & Martinez G (2004) A framework for an evolutionary path toward 4G by means of cooperation of networks. IEEE Communications 5: 82-89.
33. Aaltonen J, Grundström M, Hakulinen H, Hussman H & Talmola P (2002). Mobile Internet Technical Arcitecture 1: 201-206.
34. Kohtala A (2003) A fourth digital broadcast network - Creating a market for mobile media services in Finland. Ministry of Transport and Communications Finland Publication Series.
35. 3GPP (1999) Stage 2 Functional Specification of Location Services in UTRAN. 3GPP TSGRAN WG2 3G TS 25305 V330 (Release 1999).
36. Ko YB & Vaidya NH (1999) Geocasting in mobile ad hoc networks: location-based multicast algorithms. Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, Louisiana, USA, 1: 101-110.
37. Ko YB & Vaidya NH (1998) Location-aided routing in mobile ad hoc networks. Proc. ACM/IEEE Mobile Communications Conference, Dallas, Texas, USA, 1: 66-75.
38. Yang X, Yap JH, Ghaheri-Niri S & Tafazolli R (2001) Enhanced position assisted soft handover algorithm for UTRA. Proc. 54th Vehicular Technology Conference Fall, Atlantic City, New Jersey, USA, 2: 567-571.
39. Markopoulos A, Pissaris P, Kyriazakos S & Sykas E (2003) Optimized handover procedure based on mobile location in cellular systems. Proc. 14th IEEE Proceedings on Personal Indoor and Mobile Radio Communications, Beijing, China, 3: 2490-2494.
40. Inoue M, Mahmud K, Murakami H, Hasegawa M & Morikawa H (2004) Novel out-of-band signaling for seamless interworking betweem heterogeneous networks. IEEE Wireless Communications 2: 56-63.
41. Zhang Q, Guo C, Guo Z & Zhu W (2003) Efficient mobility management for vertical handoff between WWAN and WLAN. IEEE Communications 11: 102-108.
42. Akyildiz IF, Jiang Xie & Mohanty SA (2004) Survey of mobility management in next-generation all-IP-based wireless systems. IEEE Wireless Communications 4: 16-28.

43. Ruggaber R & Seitz J (2001) A transparent network handover for nomadic CORBA users. Proc. 21st IEEE International Conference on Distributed Computing Systems, Mesa, Arizona, USA, 1: 499-506.
44. Tourrilhes J & Carter C (2002) P-Handoff: a protocol for fine grained peer-to-peer vertical handoff. Proc. 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Lisbon, Portugal, 2: 966-971.
45. Perkins CE (2002) IP Mobility Support for IPv4. Internet Engineering Task Force RFC3344.
46. Johnson D, Perkins CE & Arkko J (2004) Mobility Support in IPv6. Internet Engineering Task Force RFC3775.
47. Omae K, Inoue M, Okajima I & Umeda N (2003) Handoff performance of mobile host and mobile router employing HMIP extension. Proc. IEEE Wireless Communications and Networking Conference, New Orleans, Louisiana, USA, 2: 1218-1223.
48. Misra A, Das S, Dutta A, McAuley A & Das SK (2002) IDMP-based fast handoffs and paging in IP-based 4G mobile networks. IEEE Communications 3: 138-145.
49. Campbell AT, Gomez J, Sanghyo Kim, Chieh-Yih Wan, Turanyi ZR & Valko AG (2002) Comparison of IP micromobility protocols. IEEE Wireless Communications 1: 72-82.
50. Chao HC & Huang CY (2003) Micro-mobility mechanism for smooth handoffs in an integrated ad-hoc and cellular IPv6 network under high-speed movement. IEEE Trans Vehic Tech 6: 1576-1593.
51. Kwon TT, Gerla M & Das S (2002) Mobility management for VoIP service: Mobile IP vs. SIP. IEEE Wireless Communications 5: 66-75.
52. Casals O, Cerda Ll, Willems G, Blondia C & Van den Wijngaert N (2003) Performance evaluation of the post-registration method a low latency handoff in MIPv4. Proc. IEEE International Conference on Communications, Anchorage, Alaska, USA, 1: 522-526.
53. Helmy AA-G, Jaseemuddin M & Bhaskara G (2004) Multicast-based mobility: a novel architecture for efficient micromobility. IEEE J Sel Areas Communic 4: 677-690.
54. Typpö V (2001) Micro-mobility within wireless ad hoc networks: towards hybrid wireless multihop networks. Diploma thesis, University of Oulu, Finland.
55. Latvakoski J, Pakkala D & Paakkonen P (2004) A communication architecture for spontaneous systems. IEEE Wireless Communications 3: 36-42.
56. Jung JW, Mudumbai R, Montgomery D & Kahng HK (2003) Performance evaluation of two layered mobility management using mobile IP and session initiation protocol. Proc. IEEE Global Telecommunications Conference, San Francisco, California, USA, 3: 1190-1194.
57. Pangalos PA & et al (2001) End-to-end SIP based real time application adaptation during unplanned vertical handovers. Proc. IEEE Global Telecommunications Conference, San Antonio, Texas, USA, 6: 3488-3493.
58. Ma L, Yu F, Leung VCM & Randhawa T (2004) A new method to support UMTS/WLAN vertical handover using SCTP. IEEE Wireless Communications 4: 44-51.
59. Henderson TR, Ahrenholz JM & Kim JH (2003) Experience with the host identity protocol for secure host mobility and multihoming. Proc. Wireless Communications and Networking Conference, New Orleans, Louisiana, USA, 3: 2120-2125.
60. Du F, Ni LM & Esfahanian AH (2002) HOPOVER: a new handoff protocol for overlay networks. Proc. IEEE International Conference on Communications, New York City, New York, USA, 5: 3234-3239.
61. Hsieh R, Zhou ZG & Seneviratne A (2003) S-MIP: a seamless handoff architecture for mobile IP. Proc. 22nd Annual Joint Conference of the IEEE Computer and Communications Societies Infocom, San Francisco, California, USA, 3: 1774-1784.
62. Politis C, Chew KA, Akhtar N, Georgiades M, Tafazolli R & Dagiuklas T (2004) Hybrid multilayer mobility management with AAA context transfer capabilities for all-IP networks. IEEE Wireless Communications 4: 76-88.
63. Matsuoka H, Yoshimura M & Ohya T (2003) A robust method for soft IP handover. IEEE Internet Computing 2: 18-24.
64. Wiljakka J (2002) Transition to IPv6 in GPRS and WCDMA mobile networks. IEEE Communications 4: 134-140.

65. Pang AC, Chen JC, Chen YK & Agrawal P (2004) Mobility and session management: UMTS vs. cdma2000. IEEE Wireless Communications 4: 30-43.
66. Grimlund O & Gudmundson B (1991) Handoff strategies in microcellular systems. Proc. 41st IEEE Vehicular Technology Conference, St Louis, Missouri, USA, 1: 505-510.
67. Wei Wu, Banerjee N, Basu K & Das SK (2003) Network assisted IP mobility support in wireless LANs. Proc. 2nd IEEE International Symposium on Network Computing and Applications, Cambridge, Massachusetts, USA, 1: 257-264.
68. Prakash R & Veeravalli VV (2000) Adaptive hard handoff algorithms. IEEE J Sel Areas Communic 11: 2456-2464.
69. Prakash R & Veeravalli VV (2003) Locally optimal soft handoff algorithms. IEEE Trans Vehic Tech 2: 347-356.
70. McNair J & Zhu F (2004) Vertical handoffs in fourth-generation multinetwork environments. IEEE Wireless Communications 3: 8-15.
71. Salamah M, Tansu F & Khalil N (2003) Buffering requirements for lossless vertical handoffs in wireless overlay networks. Proc. 57th IEEE Vehicular Technology Conference Spring, Jeju, Korea, 3: 1984-1987.
72. Soma-Reddy J & Acampora A (2003) Macrodiversity in IP based cellular networks. Proc. IEEE Wireless Communications and Networking, New Orleans, Louisiana, USA, 3: 2132-2137.
73. Sharma S, Zhu N & Chiueh TC (2004) Low-latency mobile IP handoff for infrastructure-mode wireless LANs. IEEE J Sel Areas Communic 4: 643-652.
74. Floroiu JW, Ruppelt R, Sisalem D & Voglimacci J (2003) Seamless handover in terrestrial radio access networks: a case study. IEEE Communications 11: 110-116.
75. Perkins CE & Wang K (1999) Optimized smooth handoffs in Mobile IP. Proc. IEEE International Symposium on Computers and Communications, Sharm El Sheik, Red Sea, Egypt, 1: 340-346.
76. Chung SH, Yoon H & Cho JW (2002) A fast handoff scheme for IP over Bluetooth. Proc. International Conference on Parallel Processing, Vancouver, British Columbia, Canada, 1: 51-55.
77. Wang HJ, Katz RH & Giese J (1999) Policy-enabled handoffs across heterogeneous wireless networks. Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, Louisiana, USA, 1: 51-60.
78. Sun J, Tenhunen J & Sauvola J (2003) CME: a middleware architecture for network-aware adaptive applications. Proc. 14th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Beijing, China, 1: 839-843.
79. Vanem E, Svaet S & Paint F (2003) Effects of multiple access alternatives in heterogeneous wireless networks. Proc. IEEE Wireless Communications and Networking Conference, New Orleans, Louisiana, USA, 3: 1696-1700.
80. Hautala T, Braysy T, Makela J, Lehtomaki J & Saarinen T (2003) Scalability of mobility signaling in IEEE802.11 WLAN. Proc. 58th IEEE Vehicular Technology Conference Fall, Orlando, Florida, USA, 3: 2044-2047.
81. Hautala T, Suliman I, Lehtomäki J & Saarinen T (2004) Performance evaluation of videostreaming on a heterogenous multihop mobile network. Proc. IEEE Vehicular Technology Conference Spring, Milan, Italy, 5: 2744-2747.
82. Chakravorty R, Vidales P, Patanapongpibul L, Subramanian K, Pratt I & Crowcroft J (2003) On inter-network handover performance using Mobile IPv6. Cambridge Open Mobile System (COMS) Technical Report-work done in collaboration with Laboratory of Communication Engineering .
83. M Dunmore (2003) Mobile IPv6 handovers: performance analysis and evaluation. IST-2001-32603 6Net project deliverable 32603/ULANC/DS/411/A1.
84. Alimian A & Aboba B (2004) Analysis of roaming techniques. IEEE 80211-04/0377r1, Cited October 19th 2004 from http://wwwieee802org/11/Documents/
85. Pagtiz T, Chakravorty R & Crowcroft J (2003) Proactive IP mobility management for context-aware all-IP wireless access networks. Cambridge Open Mobile System (COMS) Technical Report--Joint work with University College London.

86. ETSI (1992) Handover procedures. ETSI-GSM Technical Specification GSM 0309 Version 321 European digital cellular telecommunication system (phase 1).
87. Kim SE & Copeland JA (2003) TCP for seamless vertical handoff in hybrid mobile data networks. Proc. IEEE Global Telecommunications Conference, San Francisco, California, USA, 2: 661-665.
88. Itoh KI, Watanabe S, Shih JS & Sato T (2002) Performance of handoff algorithm based on distance and RSSI measurements. IEEE Trans Vehic Tech 6: 1460-1468.
89. Forsberg D, Malinen J, Malinen J & Kari H (2000) Increasing communication availability with signal-based mobile controlled handoffs. Proc. IP based Cellular Networks Conference, Paris, France. Cited October 19th 2004 from http://ntrg.cs.tcd.ie/htewari/papers/dynamics_ipcn2000_paper.pdf
90. Matusz P, Machan P & Wozniak J (2003) Analysis of profitability of inter-system handovers between IEEE 802.11b and UMTS. Proc. 28th IEEE International Conference on Local Computer Networks, Bonn/Königswinter, Germany 1: 210-217.
91. Akar M & Mitra U (2003) Joint power and handoff control using a hybrid systems framework. Proc. 22nd Annual Joint Conference of the IEEE Computer and Communications Societies Infocom, San Francisco, California, USA, 3: 1614-1621.
92. Tabbane S (1997) Evaluation of handover target cell determination algorithms for heterogeneous cellular environments. Proc. IEEE International Conference on Communications, Sydney, Australia, 1: 396-400.
93. Ivanov K & Spring G (1995) Mobile speed sensitive handover in a mixed cell environment. Proc. 45th IEEE Vehicular Technology Conference, Chicago, Illinois, USA, 2: 892-896.
94. Liang X, Ong FLC, Chan PML, Sheriff RE & Conforto P (2003) Mobile Internet access for high-speed trains via heterogeneous networks. Proc. 14th International Symposium on IEEE Personal Indoor and Mobile Radio Communications, Beijing, China, 1: 177-181.
95. Verdone R & Zanella A (2002) Performance of received power and traffic-driven handover algorithms in urban cellular networks. IEEE Wireless Communications 1: 60-71.
96. Wallenius E & Hämäläinen T (2002) Pricing model for 3G/4G networks. Proc. 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Lisbon, Portugal, 1: 187-191.
97. Prehofer C, Nafisi N & Wei Q (2003) A framework for context-aware handover decisions. Proc. 14th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Beijing, China, 3: 2794-2798.
98. Sanmateu F, Paint L, Morand S, Tessier P, Fouquart A & Sollund E (2002) Seamless mobility across IP networks using Mobile IP. Computer Networks 40: 181-190.
99. Zonoozi M & Dassanayake P (1997) Handover delay and hysteresis margin in microcells and macrocells. Proc. 8th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Helsinki, Finland, 2: 396-400.
100. Zonoozi M, Dassanayake P & Faulkner M (1997) Optimum hysteresis level signal averaging time and handover delay. Proc. 47th IEEE Vehicular Technology Conference, Phoenix, Arizona, USA, 1: 310-313.
101. Akar M & Mitra U (2001) Variations on optimal and suboptimal handoff control for wireless communication systems. IEEE J Sel Areas Communic 6: 1173-1185.
102. Halonen T, Romero J & Melero J (2002) GSM, GPRS and EDGE performance evolution towards 3G/UMTS. John Wiley & Sons Ltd.
103. Tripathi N (1997) Generic adaptive handoff algorithms using fuzzy logic and neural networks. PhD Thesis, Virginia Polytechnic Institute and State University.
104. Wong KD & Cox DC (2000) A pattern recognition system for handoff algorithms. IEEE J Sel Areas Communic 7: 1301-1312.
105. Hongyan B, Chen H & Lingge J (2003) Intelligent signal processing of mobility management for heterogeneous networks. Proc. International Conference on Neural Networks and Signal Processing, Nanjing PR China, 2: 1578-1581.
106. Chan PML, Sheriff RE, Hu YF, Conforto P & Tocci C (2001) Mobility management incorporating fuzzy logic for heterogeneous a IP environment. IEEE Communications 12: 42-51.

107. Wang SS, Sridhar S & Green M (2002) Adaptive soft handoff method using mobile location information. Proc. 55th Vehicular Technology Conference Spring, Birmingham, Alabama, USA, 4: 1936-1940.
108. Tripathi ND, Reed JH & van Landinoham HF (1998) Handoff in cellular systems. IEEE Personal Communications 6: 26-37.
109. Orlik PV & Rappaport SS (2001) On the handoff arrival process in cellular communications. Kluwer Academic Publishers Wireless Networks Journal 7: 147-157.
110. Orlik PV & Rappaport SS (1998) A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions. IEEE J Sel Areas Communic 5: 788-803.
111. Ruggieri M, Graziosi F & Santucci F (1998) Modeling of the handover dwell time in cellular mobile communications systems. IEEE Trans Vehic Techn 2: 489-498.
112. Wang J, Qing-An Zeng & Agrawal DP (2003) Performance analysis of a preemptive and priority reservation handoff scheme for integrated service-based wireless mobile networks. IEEE Trans Mob Comp 1: 65-75.
113. Suliman M, Pomalaza-Ráez C, Lehtomäki J & Oppermann I (2004) Radio resource allocation in heterogeneous wireless networks using co-operative games. Proc. Nordic Radio Symposium Including Finnish Wireless Communications Workshop, Oulu, Finland. (CD-ROM)
114. Roy A, Das SK & Misra A (2004) Exploiting information theory for adaptive mobility and resource management in future cellular networks. IEEE Wireless Communications 4: 59-65.
115. Gao X, Wu G & Miki T (2004) End-to-end QoS provisioning in mobile heterogeneous networks. IEEE Wireless Communications 3: 24-34.
116. Fodor G, Eriksson A & Tuoriniemi A (2003) Providing quality of service in always best connected networks. IEEE Communications 7: 154-163.
117. Cheng Y & Zhuang W (2002) DiffServ resource allocation for fast handoff in wireless mobile. IEEE Internet Communications 5: 130-136.
118. Baykal B & Akan OB (2003) A QoS-aware handoff using RSVP in next generation wireless networks. Proc. 8th IEEE International Symposium on Computers and Communication, Kemer-Antalya, Turkey, 1: 125-130.
119. Zhang B & Sun Y (2004) Cooperation of RSVP DiffServ for frequent and fast intersubnet handoff. Proc. 7th IEEE International Symposium on Parallel Architectures Algorithms and Networks, Hong Kong, China, 1: 429-432.
120. Yang X (2004) IEEE 80211E: QoS provisioning at the MAC layer. IEEE Wireless Communications 3: 72-79.
121. Zhu H, Ming Li, Chlamtac I & Prabhakaran B (2004) A survey of quality of service in IEEE 802.11 networks. IEEE Wireless Communications 4: 6-14.
122. Yile Guo & Chaskar H (2002) Class-based quality of service over air interfaces in 4G mobile networks. IEEE Communications Magazine 3: 132-137.
123. Degermark M (1999) 3GPP requirements for header compression. 3GPP TR 23922 version 100, Technical Specification Group Services and Systems Aspects, Architecture for an All IP network 1: 36-37.
124. van der Schaar M & Meehan J (2002) Robust transmission of MPEG-4 scalable video over 4G wireless networks. Proc. International Conference on Image Processing, Rochester, New York, USA, 3: 757-760.
125. Read K & Maurer F (2003) Developing mobile wireless applications. IEEE Internet Computing 1: 81-86.