DEPARTMENT OF COMPUTER SCIENCE
SERIES OF PUBLICATIONS A
REPORT A-2012-3

Access Selection Methods in Cooperative Multi-operator Environments to Improve End-user and Operator Satisfaction

Petteri Pöyhönen

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Hall 10 University Main Building, on May 2nd, 2012, at 12 o' clock.

UNIVERSITY OF HELSINKI
FINLAND

Supervisor
Lea Kutvonen, University of Helsinki, Finland

Pre-examiners
Jarmo Harju, Tampere University of Technology, Finland
Kalevi Kilkki, Aalto University, Finland

Opponent
Jouni Ikonen, Lappenranta University of Technology, Finland

Custos
Jussi Kangasharju, University of Helsinki, Finland

Contact information

Department of Computer Science
P.O. Box 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki
Finland

Email address: postmaster@cs.helsinki.fi
URL: http://www.cs.Helsinki.fi/
Telephone: +358 9 1911, telefax: +358 9 191 51120

Copyright © 2012 Petteri Pöyhönen
ISSN 1238-8645
ISBN 978-952-10-7900-9 (paperback)
ISBN 978-952-10-7901-6 (PDF)
Computing Reviews (1998) Classification: C.2.0, C.2.1, C.2.4, C.2.6
Helsinki 2012
Unigrafia Oy Yliopistopaino

Access Selection Methods in Cooperative Multi-operator Environments to Improve End-user and Operator Satisfaction

Petteri Pöyhönen

Department of Computer Science
P.O. Box 68, FI-00014 University of Helsinki, Finland
petteri.poyhonen@nsn.com

PhD Thesis, Series of Publications A, Report A-2012-3
Helsinki, May 2012, 211 pages
ISSN 1238-8645
ISBN 978-952-10-7900-9 (paperback)
ISBN 978-952-10-7901-6 (PDF)

Abstract

This dissertation introduces two new access selection strategies called the network-centric and the terminal-centric strategies. These strategies are based on a distributed access selection algorithm, which is designed to exploit the network cooperation to support handovers between access networks and operators. The algorithm development was motivated by the fact that the network cooperation hides the network boundaries and through information dissemination it enables more intelligent decision-making practices to ensure better utilization of the access resources resulting in enhanced capabilities serving end users. In addition, the network cooperation hides technology boundaries that could otherwise result in increased complexity of functionalities that are visible to the end users. A richer set of information representing both end-user and operator incentives is taken into account in the access selection process compared to how it is done nowadays. This ensures enhanced end-user experience because end users are also able to influence the access selection process by expressing their preferences such as a willingness to switch access network, which could add some delay and interrupt application sessions. For Mobile Network Operators (MNOs), this provides additional means to implement load balancing at radio cell and access network levels, which could increase the utilization rate of the network resources. In practice, this means that an operator is able to serve a higher number of users with the same resources. In order to evaluate the potential performance gains of these access selection strategies from the end-user and operator perspectives, two new performance metrics called the User Satisfaction Index (USI) and the Operator Satisfaction Indicator (OSI) are proposed. The new algorithm and its evaluation provide insights for developers and engineers on what kind of benefits a network cooperation together with an enhanced access selection could provide when supported correctly in mobile networks. For operators and equipment (handset and mobile network) manufacturers it could motivate a further awareness that access selection practices should be further developed and give indications of how this development could be done. In practice, this could mean that new standardization initiatives would be proposed in the corresponding field, for instance in the 3$^{\text{rd}}$ Generation Partnership Project (3GPP) standardization. For application developers, the algorithm and the constraints used in the decision making could help to better understand how to map application level and users' information with the information that is used in the access selection. A simulation model to represent a multi-radio access network scenario where a number of users are moving and using the access resources was designed. The model was developed especially to model how a distributed access selection algorithm in a multi-Radio Access Technology (RAT) environment including one or more operators could logically function. Using this model, a set of simulation experiments was carried out to assess the potential technical benefits resulting in the use of enhanced access selection methods. Both single and multi-operator scenarios are evaluated separately. The technical metrics for the simulation experiments were selected to measure different aspects of the access network resource usage and end users' connectivity and these metrics are a number of different types of handovers, network utilization rates based on the cell load measurements, users' disconnectivity and the ability to serve users under different network load conditions. Based on the simulation measurements, two new metrics USI and OSI are used to assess the non-technical performance of these access-selection strategies compared to the legacy strategy. The simulation results and analysis indicate that the cooperation between networks increases the network utilization, coverage and service availability when the access selection is designed to take advantage of it. For the service availability in a single operator environment, the average online time is about 10% higher when the new access selection strategies are used compared to the legacy one. For the network utilization in a multi-operator environment, the network-centric strategy results in about a 20% higher utilization rate over the legacy one when the network is not overloaded. The two new access selection strategies are also able to better benefit from the network cooperation when measured using the users' disconnectivity. For the end users, this means that they perceive better connectivity when the new access selection strategies are in use compared to the legacy one, while for the operators, this means that their network resources are utilized better, i.e., their utilization rate is higher, which implies a higher Return Of Investment (ROI). Naturally, these perceived technical benefits translate into greater satisfaction as the analysis clearly shows.

Computing Reviews (1998) Categories and Subject Descriptors:

C.2.0	Computer-Communication Networks: General
C.2.1	Computer-Communication Networks: Network Architecture and Design
C.2.4	Computer-Communication Networks: Distributed Systems
C.2.6	Computer-Communication Networks: Internetworking

General Terms:
Computer-Communication Networks, Network Architecture and Design, Distributed Systems, Internetworking, Access Selection Algorithm, Network Cooperation, User Experience, User Satisfaction Index, Operator Satisfaction Indicator

Additional Key Words and Phrases:
Simulation Experiment, Network Composition

Acknowledgements

I was working many fruitful years in the Ambient Networks project, where I had an opportunity to work with interesting and innovative people. I wish to thank all colleagues I worked with and all the co-authors for the interesting discussions and research work we did together (in alphabetical order); Ramon Aguero, Oliver Blume, Daniel Hollos, Martin Johnsson, Jan Markendahl, Kostas Pentikousis, Ove Strandberg, Haitao Tang, Janne Tuononen and Jan Werding.

My gratitude goes to Prof. Kimmo Raatikainen at the Department of Computer Science at the University of Helsinki. Professor Raatikainen was my original supervisor and he instructed me in the field of mobile communication research where I started as a new researcher. Due to his tragic passing, I was assigned a new supervisor, Adjunct Prof. Lea Kutvonen at the Department of Computer Science at the University of Helsinki. I am grateful to her for all the support and advise in my studies.

I would like to thank Adjunct Prof. Jouni Ikonen at the Lappenranta University of Technology for accepting the role of opponent at my defense and Prof. Jussi Kangasharju at the Department of Computer Science at the University of Helsinki for being custos at my dissertation.

Last but not the least, my thanks also go to my employer(s), Nokia and Nokia Siemens Networks, and all my line managers (in chronological order; Hannu Flinck, Jarno Rajahalme and Jari Lehmusvuori) for giving me an opportunity to spend some of my working hours on my studies. It did matter and helped me to put this thesis together.

Petteri Pöyhönen
Helsinki Finland, 2012

Contents

List of Figures

List of Tables

List of Acronyms

3GPP	3rd Generation Partnership Project
AAA	Authentication, Authorisation, and Accounting
AHP	Analytic Hierarchy Process
ANDSF	Access Network Discovery and Selection Function
BS	Base Station
CA	Composition Agreement
CAPEX	Capital Expenditure
DTN	Delay Tolerant Network
EDGE	Enhanced Data rates for GSM Evolution
EPC	Evolved Packet Core
FMC	Fixed Mobile Convergence
GAN	Generic Access Network
GERAN	GSM/EDGE Radio Access Network
GLL	Generic Link Layer
GPRS	General Packet Radio Service
GRA	Grey Relational Analysis
GRX	GPRS Roaming eXchange
GSMA	GSM Association
HSPA	High-Speed Packet Access

ICN	Information-Centric Networking
IEEE	The Institute of Electrical and Electronics Engineers, Inc.
IETF	Internet Engineering Task Force
ISP	Internet Service Provider
IST	Information Society Technologies
LP	Linear Programming
LTE	Long-Term Evolution
MIH	Media Independent Handover
MIP	Mixed Integer Programming
MN	Mobile Node
MNO	Mobile Network Operator
MPTCP	MultiPath Transmission Control Protocol
MRRM	Multi-Radio Resource Management
MSC	Mobile Switching Centre
OPEX	Operating Expenditure
OSI	Operator Satisfaction Indicator
PC	Personal Computer
PS	Packet Switched
QoE	Quality of Experience
QoS	Quality of Service
RA	Roaming Agreement
RAM	Random Access Memory
RAT	Radio Access Technology
RNC	Radio Network Controller
ROI	Return Of Investment

ROM	Read Only Memory
RRM	Radio Resource Management
SCTP	Stream Control Transmission Protocol
SGSN	Serving GPRS Support Node
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SMS	Short Message Service
TU	Traffic Unit
UMA	Unlicensed Mobile Access
UMTS	Universal Mobile Telecommunications System
USI	User Satisfaction Index
VoIP	Voice over IP
WiFi	Wireless Fidelity
WLAN	Wireless LAN
WWAN	Wireless Wide Area Network

Chapter 1

Introduction

Since the early days of mobile phones, these devices have become smaller in size but their computational power and available memory, both Random Access Memory (RAM) and Read Only Memory (ROM), have increased. This together with the development of new radio technologies providing a higher data transfer rates has led to the situation where the role of mobile phones has drastically changed compared to their early days. Besides conventional voice services like voice calls and Short Message Services (SMSs), smart phones are also used for other tasks that are traditionally done with desktop Personal Computers (PCs) and laptops, like web surfing, streaming media, social media, emails and gaming. Most if not all of these need some sort of data connectivity to exchange or synchronize information between mobile devices and network services and storages. This synchronization could be the retrieval of new messages, an update of user specific information such as location or status in the network service where this information could be visible to other users, or the download of a streaming video.

Some applications could be such that they need to synchronize their status periodically whereas others could be explicitly triggered by some entity. A common nominator of all these use scenarios is such that although they all need to have a network connectivity, some of them additionally need some sort of Quality of Service (QoS) like a certain amount of bandwidth or relatively short latency in order to function properly from the end user's perspective. As the number of available access networks and available applications increases, this makes the access selection more difficult compared to the situation we have today, i.e., no matter what application is in use, there is a predefined way to select access. There are some exceptions on this depending on the region and operator, like in some cases the use of Voice over IP (VoIP) might be restricted to the Wireless Fidelity (WiFi) accesses that are not operated by the Mobile Network Operator (MNO) that provides the subscription for the end user.

Traditionally the access selection in mobile networks has been done based on pre-configured and static means that mostly reflect operator incentives. However, with the standardization of the Access Network Discovery and Selection Function (ANDSF) [10] and the Evolved Packet Core (EPC) [8] [9] access selection is becoming more flexible in 3$^{\mathrm{rd}}$ Generation Partnership Project (3GPP) network architectures. For instance, in the EPC the ANDSF assists a mobile device in network discovery and selection as well in handover decisions. This function sends information to a mobile node, which then reacts to this and triggers a handover. This is an example of tight network control over access selection and mobility. Moreover, also non-3GPP accesses like The Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11 radio technologies [34] belong to the list of supported radio technologies.

This is a step in the right direction in the sense that end users have a wider set of accesses available. Moreover, this also introduces new challenges for access selection, i.e., how to choose the best available access. To answer this challenge, end users' preferences have to be taken into account in the access selection, since after all, they are the users for which accesses are provided. The distributed access selection algorithm proposed in this dissertation does that. The algorithm provides the flexibility for adjusting the importance of network and terminal constraints. So it can be adjusted to act as a tight network controlled access selection but additionally it supports sharing this control between the network and terminal.

Roaming between MNOs is typically supported at international level whereas national roaming is not so common. This is well suited for the current mobile operator business landscape, where operators are tightly vertically integrated taking care of both access provisioning and service provisioning. When the networking environment becomes more diverse both in terms of the number of deployed access technologies and the number of service and network providers then more sophisticated means for selecting an access network is needed. It needs to be ensured that this increased complexity in the business and technical landscape does not hinder the end-user experience and it is also necessary that end-user incentives and preferences are taken into account when selecting an access network.

How end-user experience is impacted by the used access selection method depends on how perceivable connection characteristics like latency and throughput are for the used application(s). So while designing new ways of doing the access selection in heterogeneous networking environment, apart for verifying their technical feasibility it is equally important to ensure that there are no drawbacks when it comes to their usability and the resulting user experience.

1.1 Problem domain

In recent years, the number of supported access technologies has been increasing as new Radio Access Technologies (RATs) have been developed and deployed. Mobile broadband has started to be an attractive alternative with the adoption of flat rate pricing. The deployment of 3rd generation access technologies in Release 5 (High-Speed Packet Access (HSPA) [11] [3] [51]) and in Release 8 (Long-Term Evolution (LTE) [4] [73]) provide an increased capacity compared to their 2nd generation counterparts such as the General Packet Radio Service (GPRS) [5] [18] and Enhanced Data rates for GSM Evolution (EDGE) [14] [95]. The 3GPP is specifying new and faster radio access technologies like LTE Advanced [2] [73] [27], which will be their first 4th generation access technology. At the same time, the deployment of short and medium range radio access technologies based on the IEEE 802.12 technologies (802.11a/b/g) has been becoming quite common, this can be seen for instance as a higher installation base of Wireless LAN (WLAN) hot spots in many public places and premises like airports, cafeterias, restaurants and offices. If this same trend continues, the mobile networking environment will become even more complex in terms of available access technologies and access providers. It is important that the end users can easily select their access network according to their communication needs.

One of the existing solutions supporting seamless handovers between 3GPP and non-3GPP radio technologies was part of the 3GPP work item called “Generic Access to A/Gb interfaces” and this work item resulted in the accepted specification for the 3GPP Release 6 ([12] [13]) and it was then renamed as the Generic Access Network (GAN). This is also known in non-3GPP contexts as Unlicensed Mobile Access (UMA). The first handsets supporting this feature were dual-mode devices having both 2nd generation radio (GSM/EDGE Radio Access Network (GERAN) [24]) together with 802.11 radio. Later on, the specifications were updated to support 3rd generation radio technologies. In the EPC, the ANDSF brings flexibility to the access selection, which could operate over access network and controlled approach where no end-user preferences are taken into account. In order to include end-user preferences to the access selection, a mapping between lower level technical parameters used in the access selection and higher level conditions and constraints representing end user’s incentives needs to be defined.

The implementation of this mapping is not straightforward and it also requires that different factors affecting the end users' satisfaction are studied to better understand what aspects make end users happy when it comes to the use of services with their mobile devices. Assuming that users have true freedom to choose the available access they want to use, then the access selection faces a multi-dimensional problem space as shown in Figure 1.1. On one hand, there are end-user incentives and operator incentives that are not always non-contradicting. On the other hand, there are technical and non-technical incentives representing different parties in a value chain that should be taken into account. Some times the information on which such a decision has to be made could be incomplete. So it is reasonable to say that generally speaking, there is no such entity as an optimal decision fully satisfying all parties under all circumstances, thus the decision making is exploiting a cooperation between networks and finding the best available trade-off between different incentives according to the current conditions and constraints. So how can we assess a decision making process in such environment? One way is to use technical metrics such as network utilization, but using such metrics does not directly show how the satisfaction of different players is affected.

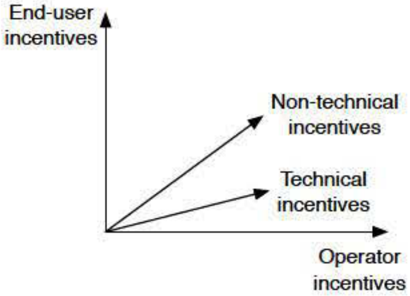


Figure 1.1: Multi-dimensional solution space

End-user incentives represent the users' communication preferences such as a QoS requirements for a new session or seamless handover support. These incentives could be derived for instance from the communication session's type, considering whether it is elastic or real time type of traffic or from other device specific settings like advanced power management settings.

Operator incentives represent involved business players in the network side such as a mobile operator or a service provider. A mobile operator incentive could be performing load balancing between radio cells or radio access networks as well as selecting the cheapest available transit provider through which user's data packets are routed.

Technical incentives represent measurements done in the terminal and network sides. There are certain limitations to doing these measurements, for example, radio signal strength measurements are clearly the measurements done in a terminal. Respectively, the current load level of Base Station (BS) and access network measurements are carried out in the network side.

Non-technical incentives are often related to some sort of monetary compensation between business entities based on their business relationship. Also, end-user preferences could be considered to be non-technical such as "save battery" or "prefer a cheaper connection" which then could be translated into different technical incentives that are considered during the access selection. For instance, a "save battery" preference could mean that an energy consumption of evaluated RATs are used as a decision making constraint.

It is important that the access selection also considers non-technical aspects since business models and related business agreements result in many kinds of operational restrictions that exist as policies and rules. In other words, even if something is technically plausible, it could be restricted by these policies and rules and therefore these business models indirectly limit the potential solution space for the access selection. There are technical challenges involved in implementing a distributed access selection algorithm, which exploits the cooperation between networks and operators. However the research presented in this dissertation does not consider the implementation aspects, thus the focus is on studying what kinds of benefits potentially an access selection algorithm using a richer set of information could result when we assume that all relevant decision making information is available. This research examines the following research questions:

- What kinds of technical benefits could a new distributed access selection algorithm exploiting network cooperation provide?
- How to model and measure the end-user satisfaction and how to measure aspects related to the operator satisfaction.
- How the gained technical benefits are translated into enhanced end-user satisfaction?

The suitability of the access selection strategies is not only evaluated in terms of technical benefits, thus also the end-user and operator satisfaction are considered, since after all these satisfaction measurements represent the real value of the gained technical benefits.

1.2 Research approach

In order to study how end-user incentives could be taken into account in the access selection and how network cooperation could help this, a new access selection algorithm is needed. The algorithm aims to exploit the network cooperation when available. Additionally, the algorithm takes into account a richer set of criteria during the decision making compared to the current practices where the access selection is done mainly based on radio measurements and under the tight control of the network. For evaluation purposes, a simulation model and implementation is required, because it is not realistic to implement and evaluate such an algorithm in a real environment. The evaluation is done both using a set of technical metrics and two new performance metrics. Technical metrics include a number of different types of handovers, network utilization rates based on the cell load measurements, users' disconnectivity and the network's ability for serving users under different network load conditions.

The evaluation compares two new access selection strategies (called the network-centric and the terminal-centric strategies) and the legacy method. All evaluated access selection strategies are implemented using the new distributed access selection algorithm. For the legacy strategy, this just means that additional decision making information is not used. Additionally, for the multi-operator environment, the handover support of the legacy strategy is limited, i.e., handovers between RATs are not supported. This was done in order to provide better reference point by which the new access selection strategies can be evaluated. The new performance index (User Satisfaction Index (USI)) models and measures the end-user satisfaction, which is one factor of the indicator (Operator Satisfaction Indicator (OSI)) representing the operator satisfaction. The performance evaluation of the algorithm and the corresponding satisfaction calculations are done based on the simulation experiments according to the developed simulation model.

The Ambient Networks project [87] [86] was a research project that was partially funded by the European Commission (Information Society Technologies (IST) 027662) and was part of the Sixth Framework Programme (FP6)[1]. There was around 40 partners including both industrial and academic partners as well as operators. The project had two phases with each phase lasting 2 years. The first phase (2004-2005) was focusing on conceptual work and the results are summarized in [52]. One of strategic goals stated for phase 1 was that the project aims to develop a new innovative network solution according to which networks could be formed

---

[1] `http://ec.europa.eu/research/fp6/index_en.cfm`

across business and technological boundaries. For multi-access, a technical goal was to develop new algorithms to manage radio resources efficiently over different technologies to enable low-cost access in multi-domain mobile networks. The second phase (2006-2007) continued the technical work based on the phase 1 results but was also focused on the validation of the developed concepts. The Multi-Access work package WP-C was developed for advanced concepts for Multi-Radio Access, however also Fixed Mobile Convergence (FMC) was considered based on the Multi-Radio Resource Management (MRRM) [84] and Generic Link Layer (GLL) [84] [70]. The work scope includes evaluation activities to assess the developed technical concepts and mechanisms. The overall Ambient Networks work including the most significant results from both phases is presented in [74] and according to the Ambient Networks vision, in the future wireless world, network accesses will be simple-to-use and they will be available anytime, anywhere and affordably for everyone. In such an environment, the access selection has naturally one of the central roles.

The developed distributed access selection algorithm is based on a high level algorithm defined in the Ambient Networks project [84] [83] [68] and additionally also extends it in order to support both terminal and network focused access selection. The new algorithm exploits a network cooperation which helps in the sharing of information between networks to ensure that the entities that are non co-located can contribute to the decision making process. The algorithm defines information specific weights based on which the access selection focus between the terminal and network constraints can be adjusted. In order to study what kinds of implications different decision making focuses have, two access selection strategies were used. The first one has its focus on the network constraints reflecting operator's incentives and preferences and is called the network-centric strategy. Correspondingly, the terminal-centric strategy has its focus on the terminal constraints representing user's incentives and preferences. The focus shift is done by using different weight sets for the access selection algorithm parameters. Unlike most of the known algorithms in this field, the new algorithm is designed to take into account higher level parameters representing end-user preferences. The simulation model was built by keeping this in mind and therefore radio modelling is simplified and the signal strength is expressed by a single value. This is the reason why the simulation results are not comparable with the existing studies where realistic radio modelling is used.

The "basic" algorithm version does not consider handover or radio bootstrapping delays. It is used to evaluate how optimal access selection decisions could be carried out compared to the theoretical upper boundaries in a single operator multi-access environment. The “advanced” algorithm version is better suited for a multi-operator environment resulting in more fine-grained access selection decisions and it also includes a handover model with radio interface and application bootstrapping delays.

For measuring what kind of impacts technical benefits have in terms of satisfaction, the user and operator satisfaction models were developed. The user satisfaction model uses and extends the approach proposed by Pohjola & Kilkki [58], where the perceived value represents the users’ happiness from the services and is used to assess the services. The user experience is divided into 3 parts; benefits, costs and success rate. Each part has own model which can be used to calculate the user experience. The proposed user satisfaction model captures each of these 3 parts in a single model that is simpler and more suitable for assessing access selection goodness. The user perception and service quality ratings are based on the findings of Tversky & Kahneman [91]. They developed a non-linear user happiness function which has a certain reference level with the function behaving differently below and above this reference level. The simulation model for the “basic” algorithm that is used in a single operator multi-access evaluation is extended to support the use of the two new performance metrics, i.e., a collection of satisfaction measurement data. For these satisfaction measurements, the “advanced” algorithm version is used and this new version is also better suited for a multi-operator environment.

Both simulation experiments use the same approach; i) perform a simulation run without the network cooperation and then ii) repeat the simulation with the network cooperation. The analysis then compares different access selection strategies without and with the cooperation as well as corresponding satisfaction calculations for the “advanced” algorithm version.

1.3 Contributions

This dissertation contains the following contributions:

- “Basic” version of the access selection algorithm, two access selection strategies (the network-centric and the terminal-centric) and the simulation model implementing these strategies in a single operator multi-access environment (Publications I and II).
- “Advanced” version of the access selection algorithm, revised access selection strategies and the modified simulation model implementing these strategies in a multi-operator multi-access environment (Publications III, IV, V, VI, VII and VIII).

- Simulation experiment results and analysis of the new access selection strategies against the legacy strategy (all publications). The simulation results and analysis show how the new access selection algorithm performs against the legacy strategy in terms of technical metrics.
- The user happiness function description, the user satisfaction index and the operator satisfaction indicator model (Publications III, IV, V, VI, VII and VIII). These publications introduce and explain the user happiness function model based on which the new performance metrics are developed.
- The user satisfaction index and the operator satisfaction indicator calculations based on the simulation experiment results and their analysis (Publications III, IV, V, VI, VII and VIII). These publications show how well the new access selection strategies performs in terms of the end-user and operator satisfaction in a multi-operator multi-access environment where the network cooperation is supported.

1.4 Structure of the dissertation

This dissertation summarizes the research work presented in the 8 original papers published in peer-reviewed scientific conferences and journals including 7 conference and 1 journal papers. The research work was mainly undertaken in the Ambient Networks project[2], which is a part of the EU's IST programme. For all the summarized publications, the author's contributions have had a central role in the development of the access selection algorithm and strategies and the simulation models. The initial ideas of creating new performance metrics was proposed by Jan Markendahl (Royal Institute of Technology)[3], and the further conceptual work in this area was a joint effort, i.e., a development of the satisfaction models and the performance metrics formulas. The author has been the main developer of the simulation model and implementation, which was originally implemented for the MATLAB®2006b program developed by MathWork™[4]. The Statistics Toolbox, which is an additional MATLAB®plug-in, was used for statistic calculations. In addition, the theoretical upper boundary calculations based on the Linear Programming (LP) technique and presented in Publications I and II were developed by Daniel Hollos (Technical University Berlin - TUB)[5] based on the provided simulation parameters

---

[2] `http://www.ambient-networks.org`

[3] email:janmar(at)kth.se

[4] `http://www.mathworks.com`

[5] email:dholloos(at)gmail.com

and settings.

- **Publication I [PI]:** P. Pöyhönen, D. Hollos, H. Tang, O. Blume, R. Aguero, K. Pentikousis, "Analysis of Load Dependency of Handover Strategies in Mobile Multiaccess Ambient Networks", Proc. of the Second Workshop on multiMedia Applications over Wireless Networks (MediaWiN 2007) and 12th IEEE Symposium on Computers and Communications (ISCC2007), July 2007, Aveiro, Portugal, pp: 15-20 [59].
- **Publication II [PII]:** D. Hollos, P. Pöyhönen, O. Strandberg, R. Aguero, K. Pentikousis, and O. Blume, "A study of handover strategies for mobile multiaccess Ambient Networks", Proc. of the 16th IST Mobile and Wireless Communications Summit, July 2007, Budapest, Hungary, pp: 1-5 [33].
- **Publication III [PIII]:** J. Markendahl, P. Pöyhönen, O. Strandberg, "Performance Metrics for Analysis of Operator Benefits of Network Cooperation in Multi-Operator Business Scenarios", The 19th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'08) [49].
- **Publication IV [PIV]:** P. Pöyhönen, J. Markendahl, O. Strandberg, "Analysis of Network Cooperation in Terms of Operator and User Satisfaction", The 7th Conference of Telecommunication, Media and Internet Techno-Economics (CTTE2008) [63].
- **Publication V [PV]:** P. Pöyhönen, J. Markendahl, O. Strandberg, "Analysis of User Experience of Access Selection in Multi-Operator Environments", The Third International Conference on Systems and Networks Communications (ICSNC2008), IEEE Computer Society. Best paper reward [64].
- **Publication VI [PVI]:** P. Pöyhönen, J. Markendahl, O. Strandberg, "Analysis of benefits of operator cooperation using end-user and operator performance metrics", 17th Biennial Conference on "The Changing Structure of the Telecommunications Industry and the New Role of Regulation", Montreal, Canada (ITS2008-Global) [62].
- **Publication VII [PVII]:** J. Markendahl, J. Werding, P. Pöyhönen, O. Strandberg, "Operator Cooperation as a Competitive Advantage for Provisioning of Low Cost High Capacity Mobile Broad band Services", 19th European Regional ITS Conference, Rome, Italy (ITS2008-Europe) [50].
- **Publication VIII [PVIII]:** P. Pöyhönen, J. Markendahl, O. Strandberg, J. Tuononen, M. Johnsson, "Analysis of Enhanced Access Selection Methods and End-User Perception in Multi-operator Envi-

ronment", International Journal On Advances in Intelligent Systems, Volume 2, pp. 107-125, 2009 [65].

The Conference of Telecommunication, Media and Internet Techno-Economics (CTTE)[6] was first time launched in 1997 and it is one major international event for the field of techno-economics. The IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)[7] was started as a workshop in 1989 and in 1992 it was first time launched as an international conference. The Intelligent Tutorial Systems (ITS) conference is an international conference that was first held in 1998 and is a major international conference for research of the ITS community[8]. The IEEE Workshop on Multimedia Applications over Wireless Networks (MediaWIN)[9] is a fairly young workshop, which was first time launched in 2006 in association with the IEEE Symposium on Computers and Communications (ISCC 2009). The IST Mobile and Wireless Communications Summit[10] is a major European event that is supported by the EU and it is an 18 years old event showcasing the on-going European research in the ICT (Information & Communication Technologies) field. The International Conference on Systems and Networks Communications (ICSNC) is a young conference arranged by the International Academy, Research and Industry Association (IARIA)[11]. The International Journal On Advances in Intelligent Systems[12] presents the selected papers from the IARIA conferences. The journal is published twice a year.

The rest of the dissertation is organized as follows; Chapter 2 explains the current network cooperation practices and presents the new developed access selection strategies based on the new distributed access selection algorithm. Chapter 3 introduces a new way to model user happiness which was used as the basis for new performance index called the USI. Chapter 4 discusses the operator satisfaction factors and related business aspects and describes another new performance metrics called the OSI to measure the operator's satisfaction. Chapter 5 illustrates the simulation models used in the simulation experiments and explains the working assumptions and other used settings. Chapter 6 shows the simulation results and satisfaction calculations and analyses them. Finally, Chapter 7 concludes the research presented in this dissertation with some final notes.

---

[6] `http://www.ctte-conference.org`

[7] `http://www.comsoc.org/confs/pimrc/index.html`

[8] `http://www.itsconference.org/`

[9] `http://mediawin.it.teithe.gr/`

[10] `http://www.ict-mobilesummit.eu/2010/`

[11] `http://www.iaria.org/`

[12] `http://www.iariajournals.org/intelligent_systems/index.html`

Chapter 2

Distributed access selection in cooperative environment

This chapter provides the examples of how network cooperation exists in today's networking environment and explains the Ambient Networks project work on the network cooperation and multi-access selection areas. The chapter describes the new distributed access selection algorithm in detail and explains what kinds of information are used in a decision making process to select the used radio access in a cooperative environment according to local and current conditions. The new algorithm and the decision making information provide insight into how the algorithm distribution could work in multi-domain environment by showing where "control loops" can exist, which could elaborate how they can be implemented in a real environment.

The decision making information presented in this chapter provides some examples on how to map user information and algorithm parameters to each other. It is not necessary to have direct 1:1 mapping between the information used in the access selection and the information visible for applications and users, but the former can be derived from the latter. For instance, if an application session is quality sensitive, then this can be used to define the access selection constraints with an aim to minimize potential connection breaks resulting from non-seamless handovers. A successful mapping naturally results in a more accurate selection process from a user's point of view.

2.1 Network cooperation

Cooperation between networks has many forms in the current environment and as the cooperation term implies, it tends to be some kind of agreement between two ('bilateral') or more ('multilateral') parties. An agreement itself normally defines responsibilities for each party according to the cooperation. One example is Service Level Agreement (SLA) between networks that represent the cooperation between parties. There are different ways to classify different types of agreements, but in this research, they are divided into horizontal and vertical agreements as illustrated in Figure 2.1. Horizontal agreements are typically between operators or between similar types of business players. Agreements have different lifetimes and an agreement between the *RAT-1* and *RAT-2* in the *operator-1* network could be interpreted as Radio Access Technologies (RATs) which are installed in "plug and play" fashion, but once installed, they become a fixed part of the operator's network infrastructure. Vertical agreements exist between different kinds of business players like a Roaming Agreement (RA) between an operator and subscription provider (also called a home operator) or an SLA between an operator and transit provider. An agreement defines how the related cooperation is going to function and what are the involved entities' responsibilities.

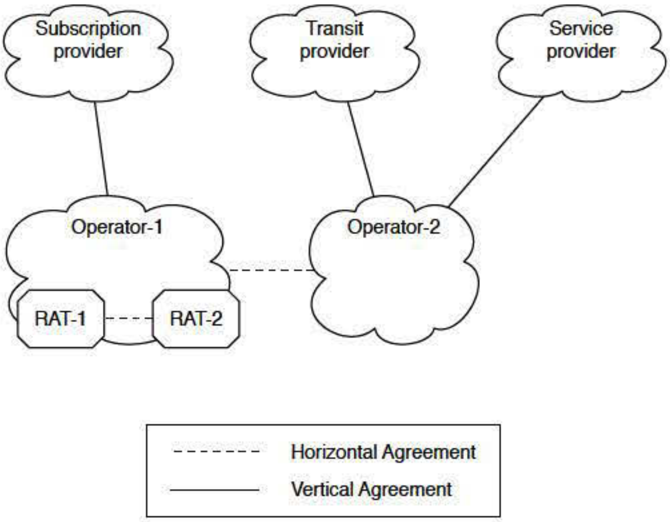


Figure 2.1: Example of different types of agreements

A cooperation could be an interworking between access networks or operator networks to support both horizontal and vertical handovers. Horizontal handovers do not involve a change of access technology, i.e., data link layer technology, and correspondingly vertical handovers require a change of access technology. Network interworking is nothing new in mobile networking area and the authors of [71] analyze means to integrate 3G and Wireless LAN (WLAN) networks. The 3^{rd} Generation Partnership Project (3GPP) had a similar scope when it specified an interworking specification between WLAN and 3GPP networks [7]. This specification defines how 3GPP subscribers can access 3GPP Packet Switched (PS) based services via a WLAN access network using the 3GPP Authentication, Authorisation, and Accounting (AAA) functions. But the specification does not specify how and when WLAN access should be preferred over 3GPP access.

Resource sharing is also one form of cooperation for instance specified for network resource sharing in [6]. According to this specification, multiple core network operators can share the same radio access network resources. There are different network configurations according to which the network sharing is configured. In one example, only Radio Network Controllers (RNCs) are shared and in another example, operators share also Mobile Switching Centres (MSCs) and Serving GPRS Support Nodes (SGSNs).

Yet another common example is national and international roaming that always requires a RA between operators to allow users to use the visited network's services. Roaming is typically carried out over a GPRS Roaming eXchange (GRX) network, which is based on agreements between operators and provides an IP backbone for 3GPP operators as specified by the GSM Association (GSMA) [31]. A number of research projects have studied network cooperation and resource control in heterogeneous networks such as [87] [88] [85].

Some form of cooperation is also required for handovers between access networks. The authors of [90] describe how handovers are implemented in cellular systems and provide a good summary of basic handover types used also in today's mobile cellular systems. There are three basic ways a handover decision can be made: network controlled, mobile assisted and mobile controlled access selection/decision making. For each type, measuring and decision making phases are distributed in a different way and each type has different characteristics like handover execution delays affecting their applicability. In the basic case, user devices send measurements to the network where a handover decision is made inside an administrative domain. This is called a mobile assisted strategy and it is used in 2G and 3G networks. This provides predictability since the network has power to make a handover decision according to the local resource management situation. This is not necessarily adequate, since things get a bit more complex in multi-access and multi-domain environments and the handover decision making including access selection gets even harder in a (competitive) multi-operator environment where competing operators do not cooperate, i.e., share information, and therefore there is no single point in the network where an optimal handover decision could be made.

The Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.21 [89] has specified Media Independent Handover (MIH) services for local and metropolitan area networks [35] allowing information gathering in a heterogeneous access networking environment to be used for access selection. The scope is mostly for vertical handovers meaning that access technology changes due to a handover, e.g., a handover between access networks. This concept helps to effectively discover and select accesses in heterogeneous environments by enabling interworking between heterogeneous link-layer technologies (e.g. cellular and IEEE 802 networks) and aims to optimize vertical handovers. For instance, the MIH supports the discovery of Wireless Fidelity (WiFi) hot spots over cellular radio. The MIH includes 3 different handover models: i) terminal controlled, ii) terminal initiated and network assisted, and iii) network initiated and network controlled. The main difference between the MIH work and the distributed access selection work of the research presented in this dissertation is that the latter also considers how to value discovered accesses in the access selection. However, this 802.21 work is a step in the right direction to support heterogeneous multi-access features. In the same vertical handover problem domain, the decision making can be roughly divided into 3 categories [78]: i) policy based decision making, ii) fuzzy logic based decision making [53], and iii) multi factor based decision making [77] [80].

The fuzzy logic based decision making [53] uses multiple factors to make a handover decision for vertical handovers between Wireless Wide Area Networks (WWANs) and WLANs. The network side is well represented in the constraints, but the user side is excluded. The handover decision aims to improve Quality of Service (QoS) by using WLAN whenever available, i.e. off-loading traffic from mobile networks to hot spots when available and feasible. Nevertheless, this is an example of how to improve handover decision making for vertical handovers using multiple decision constraints. In addition, it shows that fuzzy logic based approaches can be feasible to solve multi-constraints optimization problems in the field of access and network selection.

The multi-factor decision making approach is described in [77], where a new network selection mechanism for heterogeneous multi-access networks including both cellular and IEEE 802 radio technologies is proposed. The mechanism combines two mathematical techniques, the Analytic Hierarchy Process (AHP) and the Grey Relational Analysis (GRA). The decision making goal is to find a trade-off between user's preferences and network conditions. This results in a network selection function with a dedicated phase for data collection, which does not explicitly refer to the algorithm distribution. User preferences are mainly QoS related. The main differences are the used constraints and their arithmetic, which is also the reason why the constraints types are different, i.e., soft/hard constraints and binary/non-binary constraints. The mechanism does not exploit MIH for instance, since it requires that WLAN is discovered before its data is collected. The scope of the proposed mechanism is quite similar to distributed access selection in this dissertation; both are multi-criteria mechanism and include user's preferences. However, the distributed access selection is not limited to use only for vertical handovers, thus it is also suitable for horizontal ones, i.e., handovers inside the same access network. The relation is such that the proposed mechanism shows interesting possibilities on how potentially one could implement the distributed access selection algorithm of this dissertation, since the proposed mechanism can lead to decreased implementation complexity assuming that the decision making constraints can be expressed and computed accurately with the AHP and GRA.

In [78], enhancements are proposed for MIH framework to support vertical handovers including a network selection algorithm. This enhanced framework like the distributed access selection in this dissertation differentiate between terminal and network sides and the related decision factors. The decision making is rank aggregation based, which combines per factor ranking results to decide what is the best available network. The algorithm uses a Markov chain transition matrix with decision factor specific weights. So in this sense, it is similar to the distributed access selection in this dissertation, but the arithmetic is naturally different because the Markov chain based approach is not used in the distributed access selection algorithm. However, the proposed framework introduces a new multi-constraint decision making process and therefore provides one potential implementation environment for the distributed access selection of this dissertation.

The authors of [80] have also studied multi-constraint decision making based on a Markov decision process for vertical handover decision making where also signalling is considered. This work also considers on how to extend the original model to include the user's preferences and horizontal handover support. According to this extension, network and terminal side constraints are clearly separated (network-based v.s. user-based). For terminal side, access cost, power consumption and security are proposed as user-based constraints to be used in the decision making. This differentiation of network and terminal side constraints also follows the design approach used for the distributed access selection algorithm presented in this dissertation. Even though the used constraint sets are different in these two models, the work presented in [80] still supports one of the design decisions that was made while defining the distributed access selection algorithm.

The authors of [98] describe a multi-service vertical handover decision algorithm that does not only deal with signal strength measurements, thus it also considers other criteria like access costs and user preferences. Both mobile device and network controlled handover models are described. The optimization algorithm is based on the authors previous work [97] and unlike the distributed access selection algorithm in this dissertation, it considers user satisfaction during a hand-off decision in terms of QoS and number of service requests. The algorithm makes hand-off decisions between networks. So in one sense, user satisfaction is considered to be some kind of measurable metric that is then measured and assessed during a hand-off decision making. In other words, QoS is treated differently than in the User Satisfaction Index (USI) model, where it is relative metric and its value depends on the user expectation. The authors did simulation experiments for their algorithm but due to different simulation metrics, their results are not comparable with the results presented in this dissertation. However, their results show how more complex multi-constraint decision making can result in more effective load balancing with a relatively high network load.

The authors of [25] describe a framework to support interworking in a heterogeneous multi-access networking environment. They consider both cellular networks as well as IEEE 802 networks. The scope is different compared to the work scope described in this dissertation, since the authors also consider network layer aspects and how handovers could be done in different layers. The details do not go into any specific algorithms, thus the framework is described using architectural building blocks. Nevertheless the described work is interesting since it illustrates what kinds of architectural environments the proposed distributed access selection algorithm is designed for and shows how complex the environments for algorithm distribution and dissemination of decision making information can be.

The authors of [96] introduce a new handover decision strategy for 4th generation mobile networks where "all-IP" transport is a common nominator for all heterogeneous wireless accesses. This strategy is a mobile controlled one where a mobile node collects required information in order to perform a handover decision. In order to use this type of strategy, it is required that the network side is able to provide all the required information for the decision making. In practice, this could result in a situation where all user preferences cannot be covered sufficiently due to user's unwillingness to share personal information. This situation is even more realistic for visited users, where a user is suppose to provide personal information for the visited network.

The authors of [40] provide an overview of different vertical handover strategies for wireless heterogeneous networks and proposes a new handover strategy that combines most favorable features from existing well-known strategies. This strategy is a mobile controlled strategy, which does not require so much complexity in the network side compared to the mobile assisted strategies.

The authors of [75] propose a new vertical handover strategy that aims to optimize the whole system performance instead of focusing on any individual preferences. The proposed strategy uses two separate algorithms, one for ongoing sessions and another for new sessions. The former is basically a vertical handover decision making algorithm that is a network controlled access selection strategy and the latter is a network selection algorithm. The algorithm uses normalized cost function metrics making it easier to compare different networks. The simulation results show that this new strategy results in a lower blocking probability compared to existing vertical handover strategies.

A number of studies have done performance evaluations for vertical handovers in multi-access and multi-operator scenarios [44] [76] [32] [82]. In [44], the authors introduce a common evaluation approach for vertical handover algorithms that typically use non-compatible use scenarios making their comparison hard or even impossible. The evaluation models examined use one WWAN and one WLAN network; i.e. General Packet Radio Service (GPRS) and WiFi. Two types of algorithms are evaluated: i) hysteresis based and ii) dwelling timer based algorithms. The former is a pure location based access selection algorithm and has its problems when movement speed increases, i.e., an increased ping-pong effect. The latter is a location and time based access selection algorithm which does not suffer from the ping pong effect even with a higher movement speed. The proposed common evaluation approach uses two main metrics to make a handover decision: i) average ping-pong number indicating how often a handover is triggered during a short time period and ii) matching which reflects the QoS of a wireless medium. These metrics are different from what is used in the simulation model in this dissertation and there is no clear mapping between them making the comparison of the results impossible. In addition, the simulation model is very different for radio cell and signal strength models. The authors point out that in most research work in this field, the used movement model is such that users move at a constant speed without changing their direction. In the proposed approach users have variation in their movement patterns in terms of speed and direction. This is quite similar but not identical with the movement model used in the simulation model in this dissertation, because the maximum speed is lower (10 m/s v.s. 20 m/s) and a new randomized movement direction is limited to ±90° compared to the old direction.

The authors of [55] proposed a new access selection method called consumer surplus that is used to make an access selection in a user device, not in a network. The consumer surplus based algorithm is evaluated against an "always cheapest" selection strategy for transferring non-real time traffic using a transfer completion time metric. This algorithm has similar characteristics as one of the developed strategy in this dissertation, i.e., both emphasize the control of a user device, but still this consumer surplus based algorithm has a more narrow set of information in its decision making process compared to the developed access selection strategies in this dissertation.

There is a number of studies of how vertical handovers could be done in heterogeneous multi-access environments [46] [93]. The authors of [46] introduce a new way to execute a vertical handover between WLAN and Universal Mobile Telecommunications System (UMTS) networks using the Stream Control Transmission Protocol (SCTP), but the work does not consider the access selection and when to trigger a handover. The method exploits SCTP's multi-homing and dynamic address reconfigurability capabilities. The proposed method is verified with the simulation experiments. In a similar way, the authors of [93] introduce Session Initiation Protocol (SIP) based handovers between WWANs and WLANs. These works are mostly focusing on the handover execution phase, which complements the distributed access selection work presented in this dissertation, because the main focus is on information collection and decision making.

The Ambient Networks research project studied and developed a generic framework to dynamically establish a cooperation between any kind of network and this framework is called the network composition [38] [16]. The usability of this framework in the 3GPP networking environment and new innovative use scenarios are also studied in [1] and respectively its business implications are studied in [66].

The network composition framework describes a set of procedures that are represented in Figure 2.2 and through which networks can dynamically establish cooperation. While two networks are composing, not all these procedures are mandatory, thus some of them could be omitted and others could be repeated multiple times. For instance, if only a single radio interface is enabled, then Media Sense and Advertisement & Discovery can be performed over it and after that, additional radio interfaces could be enabled based on the received advertisements and then Media Sense and Advertisement & Discovery are performed again over these radios. Additional complexity and signaling cost for using the network composition framework in multi-operator environment is studied in [15], where the feasibility and performance of the framework is evaluated based on the simulations.

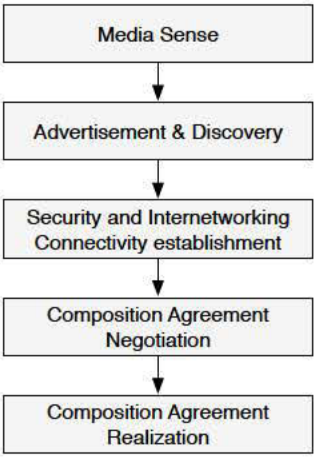
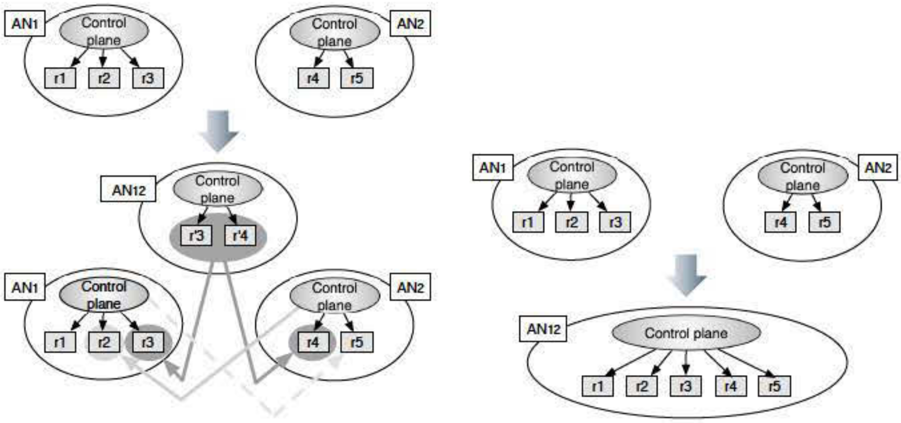


Figure 2.2: Composition procedures [16]

During the Media Sense phase, the medium for a composition is detected. The medium could be either physical or logical. The Advertisement & Discovery phase is used by networks and devices to exchange their offers and requests including capabilities, resources and different kinds of services. If adequate offers are found, then the network composition process could continue. During the Security and Internetworking Connectivity establishment phase, the composing entities such as networks establish a secure connectivity over which they can continue their composition process. It should be noted that "security level" is case sensitive and depends on the reason why two networks would like to compose. Once the "secure enough" interconnection has been established, then the involved entities start the Composition Agreement Negotiation phase during which they negotiate a Composition Agreement (CA). A CA includes details like what resources are contributed and what are the responsibilities of both parties. Additionally, a CA could also cover some commercial factors and then it is important to take care of non-repudiation, for instance by using digital signatures. Once the negotiation of a CA has been done, the Composition

Agreement Realization phase starts and the involved entities configure their local resources according to the negotiated CA. Each network configures their physical and logical resources to reflect the CA. This configuration can involve different kinds of actions depending on the content of the CA. For instance, address re-assignments could be done or admission control could be adjusted and so on.



(a) Network Interworking, Control Delegation, and Control Sharing.

(b) Network Integration.

Figure 2.3: Network Composition types [38]

There are four different composition types which the networks can use to compose themselves: i) Network Interworking, ii) Control Delegation, iii) Control Sharing and iv) Network Integration. The Network Interworking composition type does grant only access to the composed resources but does not cover control of the resources as resource $r5$ in Figure 2.3(a) where $AN1$ can access the resource, but has no control over it. The Control Delegation type includes the exclusive rights to control the composed resources. Once a network delegates the control of its resources to another network, the network loses its rights to control those resources. This is represented in Figure 2.3(a) where $AN1$ delegates the control of resource $r2$ for $AN2$. The Control Sharing enables the composing networks to share the control of resources. This is illustrated in Figure 2.3(a), where resources $r3$ and $r4$ are under shared control through a new network $AN12$ where these resources are denoted as $r'3$ and $r'4$. Finally, the Network Integration type integrates two networks into a new network, which also replaces two composing networks as presented in Figure 2.3(b). In other words, from the resource management point of view, these four composition types represent different degrees of cooperation between networks, i.e., Network Integration composition type corresponds to the strongest degree of cooperation and respectively Network Interworking represents the weakest degree of cooperation.

2.2 Access selection in Ambient Networks

There are two control plane functionalities that have a major role in the access selection in the Ambient Networks [38] [68] [84] [57] [74] type of network. The first one is the Generic Link Layer (GLL) and the second one is Multi-Radio Resource Management (MRRM). The GLL provides a generic interface to transfer information over an access link. The entities using this generic interface such as the MRRM can then be technology agnostic. The MRRM is responsible for access advertisement, access discovery and access selection functions. It also provides the service interface for other control functionalities in the control plane through which the resources of access flows can be managed. For access advertisement and discovery, conventional methods like sending and receiving radio beacons are supported.

Additionally, more advanced methods like dedicated neighborhood advertisements can be used. In case of dedicated neighborhood advertisement, a network such as an access broker can provide information about neighbor cells to the terminals. This can be used in the terminals to improve scanning efficiency and to save energy consumption. Access discovery in a heterogeneous access environment can be improved by using the networks' cell topology information in the terminals. In this way, a terminal does not always need to scan all the media, thus the scanning can be done only for the advertised accesses. Access selection is based on the concept of an MRRM set and four kinds of access sets are specified as shown in Figure 2.4:

- Detected set contains all the detected access resources by a terminal.
- Validated set contains all the access resources from the detected set that are validated by policy functions and are usable.
- Candidate set contains all access resources from the validated set satisfying the given requirements like the resource requirements of a flow.
- Active set contains the selected access resources for a flow.

In order to evaluate new access selection strategies, the algorithm needs to be further detailed. The rules by which the validated set and candidate set are constructed are not necessarily static, thus depending on the context, different kinds of policies and requirements can be used in these rules.

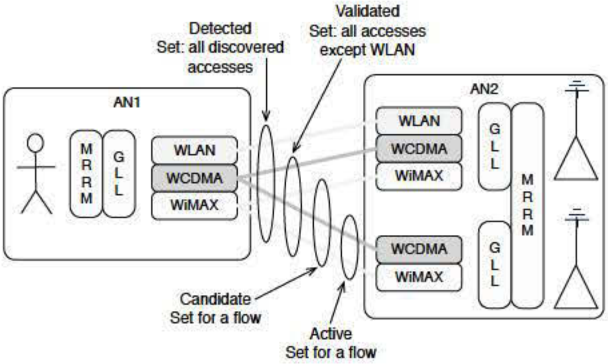
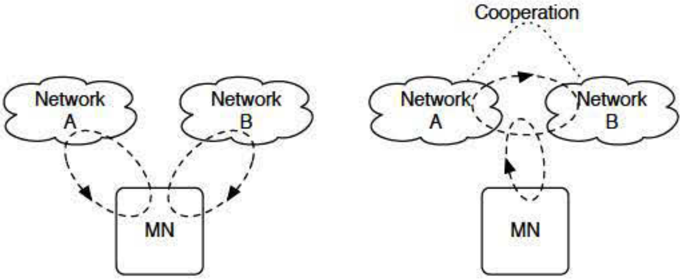


Figure 2.4: MRRM access sets

2.3 Network cooperation and algorithm distribution

Network cooperation is an enabler for the algorithm distribution, since without cooperation the entities in different networks cannot coordinate their actions directly without the assistance of a mobile node as illustrated in Figure 2.5(a). The algorithm distribution can be seen as a form of cooperation, since it requires synchronization and communication between the entities that are non co-located. Once cooperation exists between the networks, then it is possible to hide the network boundaries from a mobile node as shown in Figure 2.5(b), where Network A and B are cooperating and logically algorithm distribution only requires a single loop between a mobile node and the networks. The way the networks then implement the cooperation and the distributed algorithm is case sensitive.



(a) Without cooperation. (b) With cooperation.

Figure 2.5: Algorithm distribution and cooperation

Horizontal agreements represent a cooperation between network providers. For example, when a service continuity is preferred for an existing connection, it requires a horizontal agreement between old and new network providers. Naturally, overlapping operator coverage areas provide a good base for the cooperation between operators to support (seamless) inter-operator handovers and load balancing.

Vertical agreements are used between network and service providers. This type of agreement represents a cooperation based on which information is collected from a service provider to be taken into account in a distributed decision making process. These agreements are also required in a roaming situation where a user is accessing a visited network, which requires a valid roaming agreement with the user's subscription provider (home operator).

From the operator's perspective, both types of cooperation are important, but they have different roles under different circumstances. For instance, the ability to support inter-operator handovers might be an important criterion when a mobile terminal considers the possibility to perform a handover between operators while a "critical" application session is running. In these kinds of cases, the lack of sufficient cooperation between old and new access network might result in the disqualification of the new network's detected cell(s). Vertical agreements are more important from the access selection point of view, when a user is establishing a new connection. While considering what access to take into use, it is also important to consider what are the communication preferences.

According to the cooperation in question, providers are able to coordinate their actions, delegate or share resource management tasks and so on. A higher degree of cooperation implies wider information visibility and the degree of cooperation is related to the different network composition types described in Section 2.1. This is an important aspect especially from a distributed access selection point of view, since constraints and related conditions are distributed over administrative domain boundaries [48] [60].

A distributed access selection algorithm should be able to function on top of a diverse business landscape where technical agreements between networks and players like SLAs realize the business relationships and set operational boundaries. Algorithm distribution means that all relevant information is gathered, maintained and used according to existing horizontal and vertical agreements in the decision making.

2.4 Distributed access selection algorithm

Access selection in multi-access mobile networks is a challenging task, since all the information used in the access selection algorithm cannot be freely distributed and therefore in order to use such information algorithm distribution is required. The restrictions for information distribution could stem from business and technical agreements between the involved players limiting the information visibility. Information with a short lifetime like lower protocol stack information, for instance a radio channel conditions, or otherwise very frequently changing information can be only maintained within a limited (e.g., local) scope without that overall system performance would suffer, e.g., the control signalling for maintaining and updating the information state would take too many resources if the information changes too frequently. The second aspect favouring the distribution is the diversity of an operational environment, since the entities participating in a distributed access selection are not always co-located and therefore the communication costs between the entities vary in terms of communication delay, errors, battery consumption, etc.

An access selection strategy to find a trade-off between different preferences should be generic and flexible enough in order to support adaptability. A decision making process could be extended on-the-fly by introducing new entities and their parameters. In addition, it is even possible that the algorithm's operational mode (access selection strategy) can be changed dynamically, e.g., from a mobile controlled to a network controlled. Frequently or periodically changing information affects decision making and this can result in the need to adjust the access selection strategy, for instance to favour end-user preferences over operator ones or vice versa.

Cooperation between entities has a crucial role in a distributed access selection algorithm and it could be realized with the network composition framework. Each entity has its responsibilities and roles defined in the distributed access selection based on the cooperation agreements. Once the agreements are established, each entity knows how to autonomously operate as a part of the distributed algorithm. Cooperation between mobile operators can hide operator boundaries from the end users.

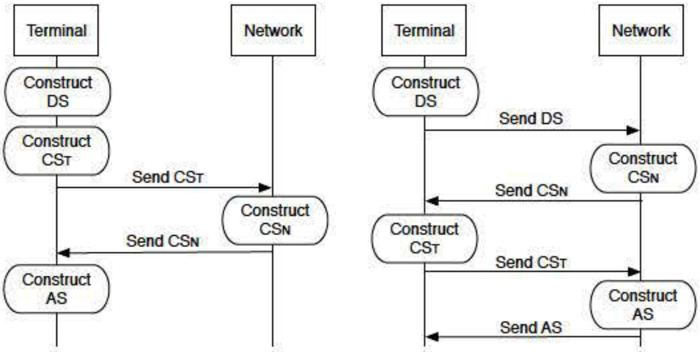
The developed algorithm uses and extends a high-level access selection algorithm described in [83] [68] with a few exceptions. First, the algorithm does not use the validated set because the simulation model does not include any special access policies and therefore this access set is not needed. Secondly, the algorithm extends the access sets by introducing two different candidate sets representing both terminal and network constraints. The same algorithm is used for horizontal and vertical handovers meaning it does not matter whether the main goal is to serve a new access request or to maintain an ongoing user session. The algorithm execution starts in a terminal based on a trigger generated either in the terminal or network. A terminal is naturally the only entity able to detect what cells are in its coverage. After this, depending on the used access selection strategy, the detected cells are further ranked based on terminal and network constraints, i.e., the order in which this ranking is done varies based on the used strategy. Finally, both the terminal and network cell ranks are considered together to decide which cell is the best one, i.e., the cell with the highest calculated rank value is selected.

Two different algorithm versions have been developed: a "basic" and an "advanced" one. The former does not include all the functional details, thus its focus is to evaluate how well an access selection can be done compared to the theoretical upper boundaries using the same metrics for both cases; average connectivity time and average offline time allowance. The latter is a more detailed version, which also considers handover execution and radio interface bootstrapping delays and is better suited for multi-operator environments. Both algorithm versions use the same access sets and access selection strategies. Their main differences are in the way they calculate cell ranks, access set usage and the used decision making constraints. These algorithms are implemented without any kind of distribution, because that is not needed for the assessment of the proposed decision making approach. Algorithms are executed in the simulation system where all necessary decision making information is available.

2.4.1 Access selection strategies

Algorithm distribution has an important role in the proposed new access selection strategies. These strategies are called the terminal-centric and network-centric strategies and they are illustrated in Figure 2.6.

In the terminal-centric case (Figure 2.6(a)), a terminal first constructs the Detected Set (DS) based on which the candidate set for the terminal (CS_T) is created. After this, the CS_T is sent to the network side where it is used to create the candidate set for the network (CS_N). Then, the CS_N is transferred to the terminal, where the final selection is done based on the two candidate sets; CS_T and CS_N. During the final selection, if the "advanced" version is used, then the cell having a highest combined rank value in both candidate sets is selected. For the "basic" algorithm, only the cells' positions in the access sets are considered and the numerical cell rank values are omitted. If two cells have the same combined rank, then in the case of the terminal-centric strategy, the cell with a higher rank in the



(a) Terminal-centric (b) Network-centric

Figure 2.6: New access selection strategies

CS_T is selected. Respectively, for the network-centric strategy, the CS_N is used. For the network-centric strategy as represented in Figure 2.6(b), the access sets are processed in different order and the final selection is done now in the network side. Otherwise, the functionality is the same as for the terminal-centric case.

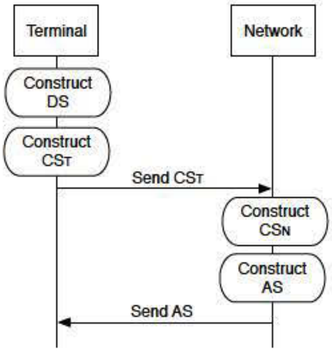


Figure 2.7: Legacy strategy

Figure 2.7 illustrates the legacy strategy that is implemented using the distributed algorithm. Comparing to the other two strategies, the main difference here is that the AS (Access Set) is done in the same place where the second candidate set (CS_N) is constructed. According to this strategy, a terminal first constructs the DS and the CS_T and sends the CS_T to the network side. Once the CS_T is received, the network side checks local conditions and constructs the CS_N and the AS, which then defines what cell the terminal should use resulting in a handover if the AS consists of a new cell that is not currently in use.

So the legacy and the network-centric strategy could be called a mobile assisted and network controlled strategy whereas the terminal-centric strategy is more like a network assisted and terminal controlled one.

Algorithm distribution implements the decision making process between network(s) and terminal and therefore it is important to have weighting between the entities contributing to the decision making. This weighting is two-level representing different 'granularities' of the decision making process. The constraint specific weights make it possible to adjust the importance levels of the information inside the entity without showing it to the outside, i.e., to other entities. In principle the constraint specific weights are access selection strategy agnostic, but they are really case sensitive. The entity specific weighting provides a means to prioritize between the entities and these weights are significant for the used strategy, because besides the overall work flow, this is another way to express the significance differences between the participating decision making entities. So from this perspective, the algorithm model should be flexible enough to respond to the challenges of the distributed access selection in a heterogeneous multi-operator environment.

However, it should be noted that even if the model is flexible enough, it still does not ensure that under all circumstances the decision making results in the optimum solution. In fact, considering the fact that users and operators might have contradicting incentives, it is impossible to find a solution that would satisfy all parties since their expectations are so contradictory. So in that sense, the decision making tries to find the best possible compromise, which could however vary depending on the strategy used implying that what is best for the terminal is not necessarily best for the network due their different and contradictory expectations and preferences. In addition, the access selection and point of attachment problems can be considered to be NP hard problems as shown in [42] [94].

2.4.2 Decision making constraints

The access selection algorithm evaluates the discovered radio cells based on their calculated rank values by the configured decision making constraints. All constraints are classified according to the following two distinct factors: i) based on the value type of a constraint (binary constraint vs. non-binary constraint) and ii) based on how the constraint's conditions should be satisfied (hard constraint vs. soft constraint). These constraints basically define what aspects are considered in the decision making phase. Also, the constraint specific weighting provides flexibility in the sense that the same algorithm can be used in different cases where the decision making considers the importance of the used information differently. The algorithm can be adjusted by selecting an appropriate set of constraint and weighting them properly. What makes an appropriate choice is really case sensitive and there is no generic rules on how to do it. However, there are some mandatory constraints like signal strength inherent in the used technologies. Similarly, due to the business landscape and practices, in a multi-operator environment roaming agreement is another needed constraint.

Next, a simple example without using the real values to illustrates how the *CellRank* vectors are constructed and used is presented. Let us assume that there are 2 cells (a, b) in the *DetectedSet* and that there are 2 terminal constraints (A, B) and network constraints (C, D). First, both the terminal $(tc_A = \lfloor tc_{A,a}, tc_{A,b} \rfloor, tc_B = \lfloor tc_{B,a}, tc_{B,b} \rfloor)$ and network $(nc_C = \lfloor nc_{C,a}, nc_{C,b} \rfloor, nc_D = \lfloor nc_{D,a}, nc_{D,b} \rfloor)$ constraint vectors are constructed. After this, the constraint vectors are normalized, i.e., a vector element value is between 0 and 1, and multiplied by the constraint specific weights and summed together resulting in $CandidateSet$. For the terminal, the candidate set is calculated as:

$$CS_T = [\sum_{i=1}^{2} \lambda_i tc_{i,a}, \sum_{i=1}^{2} \lambda_i tc_{i,b}]. \tag{2.1}$$

Respectively the candidate set for network is calculated as:

$$CS_N = [\sum_{i=1}^{2} \kappa_i nc_{i,a}, \sum_{i=1}^{2} \kappa_i nc_{i,b}]. \tag{2.2}$$

Next, both $CandidateSets$ (Equation (2.1) and Equation (2.2)) are multiplied by the terminal algorithm (α) and the network algorithm (β) weights and summed together resulting in the *CellRanks* vector consisting of two elements, one for cell a and another one for cell b. The rank value for cell a is calculated as:

$$CR_a = \sum_{j=1}^{2} \alpha \lambda_j tc_{j,a} + \sum_{j=1}^{2} \beta \kappa_j nc_{j,a}, \tag{2.3}$$

and for cell b as:

$$CR_b = \sum_{j=1}^{2} \alpha\lambda_j tc_{j,b} + \sum_{j=1}^{2} \beta\kappa_j nc_{j,b}. \tag{2.4}$$

The *ActiveSet* is then constructed based on these cell rank values (Equation (2.3) and Equation (2.4)) as follows:

$$AS = \max(CR_a, CR_b), \tag{2.5}$$

where the cell with a higher rank value forms the *ActiveSet*.

2.4.3 Basic algorithm

In the "basic" algorithm, the entities participating in the decision making process located in a terminal and network communicate by exchanging the access sets. The order in which access sets are exchanged and processed is defined by the used access selection strategy. Each exchanged access set except the Detected Set (DS) is ordered based on the calculated cell ranks and only cell IDs are exchanged, not their calculated cell rank values based on which they were ordered in the first place. Also, if a cell has rank value 0, then the cell is omitted from the access set. The first constructed access set is always the DS. As the decision making process proceeds, the number of cells in an access set is potentially getting smaller; $AS \subseteq CS_2 \subseteq CS_1 \subseteq DS$ (see Figure 2.8). Only the cells that exist in both Candidate Sets (CS_1 and CS_2) can be selected to be in the active set. Potentially, there are multiple cells in the active set as in the case of multi-homing, where a mobile node is connected over multiple interfaces at the same time.

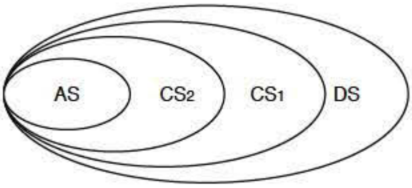


Figure 2.8: Relations between the access sets

In a terminal and network, the cell rank values are calculated for each cell in the received access set as:

$$CR_i = \sum_{j=1}^{K} \gamma_j c_j, \tag{2.6}$$

where the number of local constraints is K and γ is a constraint specific weight based on which the constraints can have different significance levels. Depending on the location, constraints are either terminal constraints or network constraints. Once the cell ranks are calculated, then the cell IDs are sorted accordingly, i.e., the cell with the highest rank is the first vector element and so on. During the final selection phase, the cell with the highest combined element position is selected. If two or more cells have the same combined position, then a terminal or network depending on which one is performing the final selection can decide according to its own constraints. A cell rank calculation could result in 0 value, if for instance the processed cell does not satisfy a hard decision making constraint (see Section 2.4.2).

Figure 2.9 shows a high level pseudo code of the "basic" access selection algorithm version, which is executed once per time unit. The algorithm gets a set of mobile nodes as its input parameter. After this, the order in which the mobile nodes are processed is randomized. For each mobile node, first the *DetectedSet* is constructed. This set contains all the radio cells a mobile node can detect and which have enough available resources according to the mobile node's demands. The resource demands are derived from an application session which is either running or which will be initiated once a new access is attached.

Once the *DetectedSet* is created, the *ConstraintSet*1 is constructed for cell rank calculations and then each cell in the *DetectedSet* is evaluated according to Equation (2.6) and the resulting cell rank value is stored to the *CellRanks* vector. Once all cells are processed, then the *CandidateSet*1 is created and it represents either CS_T or CS_N depending on the used access selection strategy. After this, the *ConstraintSet*2 is constructed for cell rank calculations done for each cell in the *CandidateSet*1. Based on these calculations, the *CandidateSet*2 is then constructed and the best cell is then searched using the two candidate sets. If the best cell is not the one, which the mobile node is currently using, then a handover is performed and if there was no errors during the handover, then the *ActiveSet* is updated with the new cell information.

This algorithm has its disadvantages in a multi-operator environment, since the way the access sets are used as ordered lists is not adequate to express how much worse n^{th} cell is compared to $(n-1)^{th}$ cell. For instance, lets consider two ordered access sets (A and B) with only cell

```
Input: Set of mobile nodes
Output: Error status
randomize the order of mobile nodes;
foreach Mobile node i do
    Read current mobile node status;
    Update mobile nodes location info;
    Construct DetectedSet;
    Construct ConstraintSet1 according to the used strategy;
    foreach Cell j in the DetectedSet do
        Calculate cell rank value using ConstraintSet1;
        Store the rank value to the CellRanks;
    end
    CandidateSet1 = ConstructSet(DetectedSet,CellRanks);
    Reset CellRanks;
    Construct ConstraintSet2 according to the used strategy;
    foreach Cell j in the CandidateSet1 do
        Calculate cell rank value using ConstraintSet2;
        Store the rank value to the CellRanks;
    end
    CandidateSet2 = ConstructSet(CandidateSet1,CellRanks);
    BestCell = FindMaxPosition(CandidateSet1,CandidateSet2);
    if CurrentCell != BestCell then
        Perform handover;
        Update the ActiveSet;
    end
end
```

Figure 2.9: Basic algorithm for evaluating cells [PI] [PII]

IDs; $A = [2, 14, 5, 7]$ and $B = [5, 2, 14, 7]$. This results in the selection of the cell with ID=2, since it has the highest combined order number: the 1st place in A and the 2nd place in B. Let us assume that the respective cell rank vectors based on which the cells are ordered (A' and B') are as follows; $A' = [3.2, 3.13, 3.1, 2.8]$ and $B' = [4.2, 3.1, 3.05, 2.98]$. The situation is now different, since the sum of the cell ID 2's rank is 6.3 (=3.2+3.1) and respectively for the cell 5 it is 7.3 (=3.1+4.2), and therefore cell ID 5 should be the most appropriate cell. This “flaw” is fixed in the “advanced” algorithm version that is introduced in Section 2.4.4.

It is possible to distribute this algorithm to support the different access selection strategies illustrated in Figure 2.6 and Figure 2.7. However, it should be noted that the decision making logic of this algorithm sets restrictions on how algorithm distribution can be done, because the logic is fairly linear and does not allow parallelism. This succession is clearly visible in Figure 2.9, where the processing order of the access sets is fixed. In other words, a creation of the $CandidateSet1$ requires that the $DetectedSet$ is already constructed and respectively the $CandidateSet1$ has to be created before the $CandidateSet2$ is constructed, etc. The distribution would follow the access strategies (Figure 2.6 and Figure 2.7), where the involved entities are aware of their local constraints and only the ordered access sets are exchanged between the entities.

2.4.4 Advanced algorithm

The "advanced" access selection algorithm that further enhances the "basic" algorithm is presented in Figure 2.10. The algorithm gets a set of mobile nodes as its input parameter. After this, the order in which the mobile nodes are processed is randomized. For each mobile node, the $DetectedSet$ is first constructed. This set contains all the radio cells a mobile node can detect and which have enough available resources according to the mobile node's demands. The resource demands are derived from an application session which is either running or which will be initiated once a new access is attached. Once the $DetectedSet$ is done, then both the network ($NetworkConstraints$) and terminal ($TerminalConstraints$) constraints are constructed. After this, each cell in the $DetectedSet$ is evaluated according to Equation (2.7) and the resulting cell rank value is added to the $CellRanks$ vector. If the cell with the highest rank is not the one which the mobile node is currently using, then a handover is performed and if there was no errors during the handover, then the $ActiveSet$ is updated with the new cell information.

Compared to the "basic" algorithm version, the "advanced" version uses the calculated cell rank values and therefore is better suited for use in a multi-operator environment. Equation (2.7) shows how the cell rank value CR_i is calculated for cell i,

$$CR_i = \alpha CRt_i + \beta CRn_i, \tag{2.7}$$

where α is the weight for the terminal constraints CRt_i and respectively β is the weight for the network constraints CRn_i. The terminal constraints CRt_i are calculated as shown in Equation (2.8):

```
Input: Set of mobile nodes
Output: Error status
randomize the order of mobile nodes;
foreach Mobile node i do
    Read current mobile node status;
    Update mobile nodes location info;
    Construct DetectedSet;
    Construct NetworkConstraints;
    Construct TerminalConstraints;
    foreach Cell j in the DetectedSet do
        Calculate cell rank value using NetworkConstraints;
        Add the rank value to the CellRanks;
        Calculate cell rank value using TerminalConstraints;
        Update the rank value to the CellRanks;
    end
    BestCell = CellWithMaxRank(CellRanks);
    if CurrentCell != BestCell then
        Perform handover;
        Update the ActiveSet;
    end
end
```

Figure 2.10: Advanced algorithm for evaluating cells [PVIII]

$$CRt_i = \sum_{j=1}^{N} \lambda_j tc_j, \tag{2.8}$$

and where λ is a constraint specific weight for the terminal constraints. Correspondingly, Equation (2.9) shows how the network constraints CRn_i with a constraint specific weight κ are calculated:

$$CRn_i = \sum_{j=1}^{M} \kappa_j nc_j. \tag{2.9}$$

If two of more cells have the same CR_i, then depending on where the final selection is done, the corresponding unweighted partial cell rank can be used to select the best cell; CRt_i or CRn_i. The partial algorithm weights α

and β are adjusted based on the used access selection strategy and the used values are shown in Table 2.1. For the network-centric strategy $\alpha > \beta$ and correspondingly for the terminal-centric one $\beta > \alpha$. For the legacy case, the same weight value was used for both algorithm weights. $\{1, 3\}$ weight value pair was selected to emphasize the differences between strategies.

Strategy Name	α	β
Terminal	3	1
Network	1	3
Legacy	1	1

Table 2.1: Algorithm weights

This algorithm is better suited for distribution than the “basic” one, because it supports some degree of parallelism. Once the *DetectedSet* is constructed and distributed, then both terminal and network side can calculate their *CellRanks* independently on each other as shown in Figure 2.10. After this, one of the calculated *CellRanks* is sent to the entity that is responsible of constructing the *ActiveSet* (this entity is defined by the used strategy as illustrated in Figure 2.6 and Figure 2.7). So like in the “basic” algorithm case, all involved entities know their local constraints and only one of the calculated *CellRanks* is exchanged between the entities. Additionally, the *AccessSet* needs to be transferred between entities if the used access selection strategy requires so.

Chapter 3

User experience model

There are many ways to measure users' happiness including many different technical aspects. When considering mobile users some additional metrics are relevant compared to the stationary users with non-mobile devices such as battery lifetime, roaming and service agreements implying where a user can have a network connection and what services are usable. Maybe one of the most traditional ways to measure user's happiness is bound to Quality of Service (QoS). QoS itself should be considered as a relative metric due to the different characteristics and needs of applications. For instance, some application types are more delay sensitive than others, and then there are Delay Tolerant Network (DTN) types of applications that still work well even if the end user experiences some random connection breaks.

Another aspect is pricing, if a user pays more for connection or service, then the user naturally expects to get a better QoS than another person paying less for the same connection or service depending on the pricing scheme used. An understanding of how the user happiness model functions and how the model is built could help us to better understand the challenges of developing the access selection mechanisms that satisfy both Mobile Network Operator (MNO) and end users. This is not only a technical engineering challenge, thus it is important to understand at least to some degree the related non-technical aspects. The User Satisfaction Index (USI) bridges (non-technical) user happiness and (technical) access selection and could help us to understand how user incentives could be expressed and used in a technical implementation of the access selection and decision making and what aspects would be beneficial to be taken into account from the user experience perspective.

3.1 Modeling of user happiness

The work presented in [28] analyzes how more sophisticated Radio Resource Management (RRM) policies affect user satisfaction in terms of transmission delay and throughput for multimedia traffic. The results show clearly that the multi-channel assignment technique increases the user satisfaction due to lower delays and higher throughput.

Badia et al. [17] model user satisfaction taking into account requested QoS, data rate and price. This model enables analysis of the impact of resource allocation on operator revenues. The model also recognizes that user behaviour depends on the used pricing. From the operator perspective, a challenging task is to find an attractive and competitive pricing model, which still ensures adequate revenues.

Lindemann et al. [43] propose an approach to increase quality of service for 3G mobile subscribers and network and service provider revenues based on real-time monitoring of the number of active mobiles and the packet loss probability. This work shows improvements for end users in terms of technical QoS measurements resulted in lower packet loss and handover failure probabilities. It also shows that when end-user satisfaction is increased, this results in better operator revenue. However, it should be noted that the revenue increase depends on the pricing model used.

Edell & Varayia [22] [23] present market and technology trial results on how users value different QoS (rate) for fixed Internet access. In this trial, users were able to dynamically adjust their QoS requirements and it was examined how users responded to different prices. One of the main findings is that the demand is sensitive to both quality and price. This work also considers different pricing aspects and the risk of “waste” of resources when flat rate pricing is used.

The authors of [72] [92] introduce a new approach to model user satisfaction which challenges the traditional customer satisfaction models basing on the assumption that “More is better”. In this Kano model, a linear relationship between different attributes is not assumed and therefore the model does not claim that for instance if a bandwidth is increased 10 times, also the customer satisfaction will increase to the same extent. According to the Kano model, certain attribute categories have a higher correlation with customer satisfaction than others. Additionally, customers value different attributes. The Kano model defines three categories of requirements and each category influences customer satisfaction in different ways; i) “Must-be” requirements, ii) “One-dimensional” requirements and iii) “Attractive” requirements. Figure 3.1 shows how these three types of requirements are related to customer satisfaction.

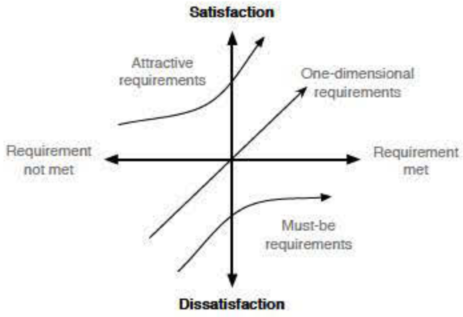


Figure 3.1: Customer satisfaction in the Kano model

"Must-be" requirements are considered as mandatory and self-evident requirements and if all these are fulfilled, then this leads only to the sate where a user is not dissatisfied. "One-dimensional" requirements linearly correlate with the requirement fulfillment level. The better a requirement is met, the more customer satisfaction is increased. "Attractive" requirements have the greatest positive impact on customer satisfaction. These requirements are not explicitly expressed, but if these requirements are met, then customer satisfaction is increased more than proportionally. However, if these requirements are not met, then this does not lead to a feeling of dissatisfaction.

There are so many ways to model user happiness and each model has its own scope and metrics based on which the happiness is then measured. Most of these models use some form of QoS in measuring user happiness. This is also a valid starting point for user happiness modeling from the access selection point of view. However, none of these existing models is suitable as such for user happiness evaluation in the access selection area, thus a need for a new measure for user satisfaction. This function needs to being able to express the user happiness differences using the measurements from the simulation experiments.

3.2 User happiness function

The user satisfaction model uses and extends the approach proposed by Pohjola & Kilkki [58], where the perceived value represents users' happiness they receive from the services and is used to value services. The user experience is divided into 3 parts; benefits, costs and success rate. Each part has its own model which together act as bases for calculating the user experience. The proposed user satisfaction model captures each of these 3 parts into a single model that is simpler and more suitable for assessing the access selection goodness. The USI does not consider absolute QoS thus it relies on the proportional concept of the service quality level. In other words, this means that a delay sensitive real-time service perceiving low delay and high bandwidth connectivity satisfying its needs are considered to be the same measured service level as in the case of a web based service perceiving a longer delay and low bandwidth connectivity also satisfying the application needs. The authors of [58] describe how to use the Perceived Experience and the Expected Value functions to represent user happiness and this is represented in Figure 3.2. In the figure, the user experience axis indicates how happy a user is and naturally its value range depends on how user experience is calculated. For the user happiness function, the value range of Y-axis and its exact numerical values are not important, thus it relies on 4 points that are marked in Y-axis as *alpha*, *beta*, *chi* and *delta*. These points represent different perception levels.

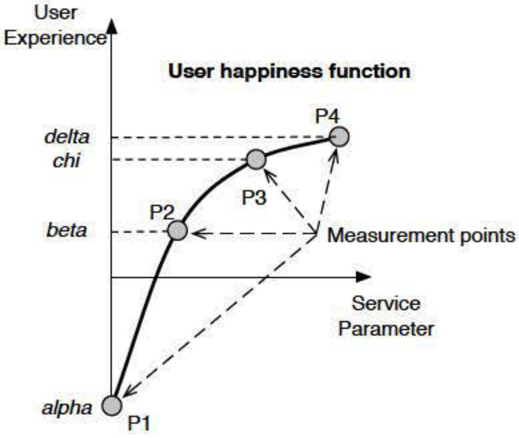


Figure 3.2: Measuring user happiness

The selected modelling approach makes two working assumptions: i) users have some level of expectation about the service availability and quality and ii) the impact of "no service" like disconnections needs to be included. The user perception and service quality ratings are based on the findings of Tversky & Kahneman [91] resulting in the non-linear user happiness function which has a certain reference level and the function behaves differently below and above this reference level. Disconnections in the model are considered as a negative value of the experience. A zero level for the user experience axis means just that the user "does not care". The simulation model for the "basic" algorithm that is used in a single operator multi-access evaluation is extended to support the use of the two new performance metrics, i.e., a collection of satisfaction measurement data. For these satisfaction measurements, the "advanced" algorithm version is used and this new version is also better suited to a multi-operator environment.

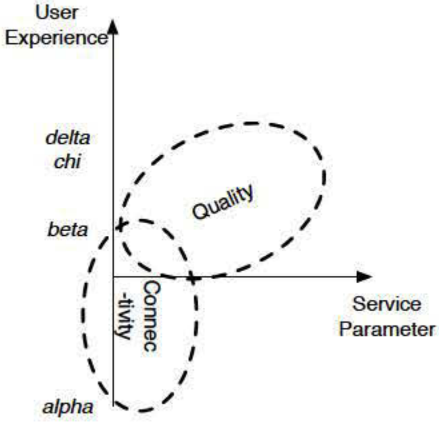


Figure 3.3: Mapping between user experience and connectivity and quality sensitivity

The USI model [PIII] [PIV] [PV] [PVI] [PVII] [PVIII] is based on a user happiness function consisting of 4 perception levels. These levels represent different perceived service qualities using different P classes; $P1$, $P2$, $P3$ and $P4$, see Figure 3.2. The class $P3$ represents the user happiness when the service parameter has the expected value. The $P2$ and $P4$ classes represent the user experience when the service parameter is lower or higher than the expected values. The $P1$ class illustrates the situation with no service due to a connection break for instance. As Figure 3.2 illustrates, the better the expectations are satisfied, the happier the user will be. In order to adjust the importance of each 4 levels, the following P class weights are needed: *alpha*, *beta*, *chi* and *delta*. Figure 3.3 shows how connectivity and quality sensitivity can be mapped to the user experience and what are the related P class weights.

Thus $P3$ represents the case where the perceived value is equal to the expected value. If the delivered service quality level is lower or higher than the expected value ("norm"), then the perceived value is worse (*beta*) or better (*delta*) than the adequate perception (*chi*). The typical perception of a person is reflected with non-linearity of the user happiness function where the slope of the user happiness function is lower when the value of expected value is increasing ($chi - beta > delta - chi$) [91].

Four levels are selected based on the assumption that a user always has a certain expectation level which can be used to evaluate the perceived service quality. This expectation level can vary based on various things like service type, user's earlier experiences and so on. This assumption alone leads to 3 different levels and is adequate as such to model the quality sensitivity. However, in order also to model the connectivity sensitivity, disconnections during which there is no perceived service needs to be differentiated from the situations where a user perceives the service quality to be worse than what was expected. So in order to model both the connectivity and the quality sensitivity, 4 different levels are required. The model could be extended by introducing more measurement points resulting in more fine grained modelling of perceived service qualities, but 4 levels are sufficient for evaluation purposes. It should be noted that some (non-elastic) services might not tolerate a large range of fluctuation in service quality level resulting in an unusable service if the perceived service quality corresponds only to $P2$. This can be captured in this model by lowering the $P2$ class weight.

3.3 User satisfaction index

The USI for a single user based on 4 different P classes with adjustable P class weights can be expressed as:

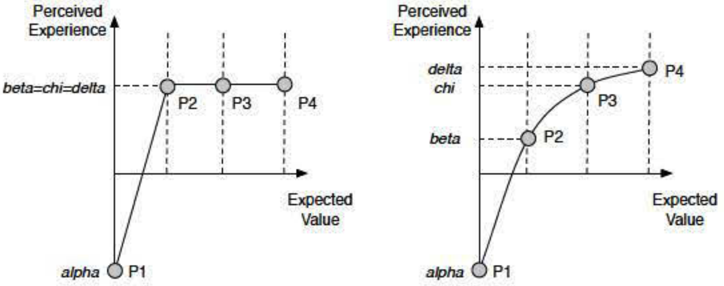
$$USI = \sum_{j=1}^{K} (\alpha * X + \beta * Y/RCost_j + \chi * Z/RCost_j + \delta * W/RCost_j), \tag{3.1}$$

where X, Y, Z, W are, respectively, the numbers representing the occurrences of the perceived service qualities $P1$, $P2$, $P3$, $P4$ in the measurements, K is the number of services, $\alpha, \beta, \chi, \delta$ are P class weights that are common for all services and $RCost$ represents relative cost associated to a transferred data unit. In addition, it assumed that there is always some relative cost associated to the data transfer, even in the case of flat rate subscription. The relative cost does not directly express how much a user has to pay for data transfer, i.e., it is not a monetary unit. When multiple operators provide the same service, then the USI should be calculated separately since the relative cost reflects the operator specific pricing and the user's overall USI consists of the USIs from each studied operator. The overall USI for multiple users is calculated as:

$$USI_{all} = \sum_{i=1}^{L} USI_i. \tag{3.2}$$

However, it is not a straightforward task to select what weights to be used. One important thing to be considered is that how much negative impact a disconnection should have in the USI calculations. Disconnection could result from many reasons like a user request for an unsupported service or an operator not having enough available access resources to serve a user. So the $P1$ class representing the disconnections captures both cases with the lack of access resources and the lack of vertical agreements between an operator and a service provider.

Figure 3.4 illustrates how different user happiness functions can be modelled by adjusting these P class weights. In Figure 3.4(b), the user happiness function is modelled using 4 different weights; $alpha = -1$, $beta = 0.25$, $chi = 1$ and $delta = 1.2$, and this corresponds the connectivity and quality sensitive case. Respectively, the connectivity sensitivity can be modelled using only two different weight values (-1 for *alpha* and 1 for the rest) as represented in Figure 3.4(a).



(a) Connection sensitive.

(b) Connection and Quality sensitive.

Figure 3.4: User happiness function examples

In both cases, the norm perceived experience ($P3$) uses the weight value 1. From simulation experiments point of view, it is important to consider both cases represented in Figure 3.4, because in these two cases, the gained technical benefits translate differently into enhanced satisfaction and one of the research question is to study how the gained technical benefits translate into enhanced satisfaction for the end users.

Chapter 4

Operator satisfaction model

The measurement of operator satisfaction is not as straightforward a procedure as measuring end-user satisfaction, because when considering operator satisfaction non-technical aspects relating to the business and marketing aspects should not be forgotten. Operator satisfaction can be estimated from many different perspectives. For instance, from a monetary perspective, the revenue an operator makes can be measured. However the revenue is not a self-contained metric, thus it correlates with many other metrics like the size of the operator's customer base, which is not constant due to churn as also implied by [41]. The size of operator's customer base changes over time for example depending on the business landscape's stability and diversity (competitors, pricing and value added services) and the users' happiness (unhappy users are more likely to switch operator if they have a choice). Additionally, the size of customer base plays an important role in the revenue formula.

Another thing to be considered is roaming users, the better roaming agreements an operator has the better abilities the operator has to serve visiting users generating extra revenue in the form of roaming fees. The Operator Satisfaction Indicator (OSI) clearly shows that there is a correlation between end-users satisfaction and operator satisfaction. Unlike the User Satisfaction Index (USI), the OSI is highly related to the economical and monetary aspects and compared to the USI it provides a new perspective from where to consider the access selection. This could help to better understand how the pricing model used affects access selection decision making.

4.1 Business landscape for cooperation

The modelling of the operator's satisfaction also requires consideration of monetary aspects. This is not a simple thing to do, since typically all agreements between business parties involving a monetary compensation are confidential. Therefore, the presented OSI model is a high level model consisting of basic monetary components like negative and positive cash flows based on the current mobile operator and Internet business landscape. However, as the authors of [47] show, in future, there might be a need for new means of compensation to support new kinds of business cases. The presented OSI model [PIII] [PIV] [PV] [PVI] [PVII] [PVIII] is based on the "conventional" mobile operator business. In addition, it should be noted that the original OSI model used the Index term, i.e., Operator Satisfaction Index, which was misleading since the OSI was not developed to be a scalar variable based on which different simulation results can be unequivocally ordered, and therefore the Index term was replaced by an Indicator term in this dissertation.

From an operator perspective, it is important to consider two monetary aspects: i) revenues and ii) costs. The former is positive cash flow and the latter is negative. So when revenues are higher than costs, an operator makes a profit. It is also important to consider how happy an operator's customers are and how well an operator is able to maintain its customer base, since perhaps the main form of revenue comes from subscribers.

A mobile operator gets revenue from different kinds of services its customers use such as voice and data services. The amount of revenue an operator makes depends on the used charging model and volume. The usage based charging model is where a user is charged based on transferred data. Flat rate charging has become more common and is mostly used with packet data services. So basically, the more customers an operator has, the more revenue potential there is depending on the charging model(s) used. Different charging models and their variations are further discussed in Section 4.3.

On the other hand, when a user uses a service, there is always some costs generated for an operator. How much cost a service usage generates, depends on many things. For instance, this cost could be Radio Access Technology (RAT) and network architecture specific relating to the deployment investments like Capital Expenditure (CAPEX) or to the operational costs like Operating Expenditure (OPEX).

4.2 Operator Satisfaction Indicator

The OSI can be represented as a value pair of the overall USI (USI_{all}) and the cost revenue function (f_{RC}) including network costs and revenue aspects:

$$OSI = \{USI_{all}, f_{RC}(Revenue_{all}, Cost_{all})\}. \tag{4.1}$$

These two values are not comparable with each other although they do correlate. For example, if the OSI_1 is $\{n, m\}$ and the OSI_2 is $\{l, k\}$ then assuming that $n < l$ and $m > k$, it cannot be concluded that the OSI_2 is better than the OSI_1. These values represent different operator's satisfaction dimensions corresponding to different metrics which can be used to measure the satisfaction. The overall USI measures how happy an operator's customers are while the cost revenue function is a monetary metric.

Another way to think about the correlation between the overall USI (USI_{all}) and the cost revenue function (f_{RC}) is to consider the USI_{all} as a trend indicating how the f_{RC} will change in future. However, how long it will take until the f_{RC} reacts to the overall USI changes is case sensitive. This is the result of how happier users potentially (in the mid or long term) result in a larger customer base and more customers results in potentially increased operator's revenue depending on the pricing used. However, it should be noted that there are many other influencing factors like market situation, competition and the business models involved.

The cost revenue function (f_{RC}) is a metric where $Revenue_{all}$ refers to the revenues generated by the users during a certain time period and respectively $Cost_{all}$ refers to the operator's cost associated with the generated traffic by the users during a certain time period. The generic form of the cost revenue function can be represented as:

$$f_{RC} = Revenue_{all} - Cost_{all}. \tag{4.2}$$

For the simulation experiments where static cases are analyzed, usage based fixed pricing is used where the price of each transferred data unit is defined based on the service type used. According to this, the cost revenue function (f_{RC}) for a single user j for a number of time intervals T_i can be expressed as:

$$f_{RC} = \sum_i D_i * P_i - D_i * C_i. \tag{4.3}$$

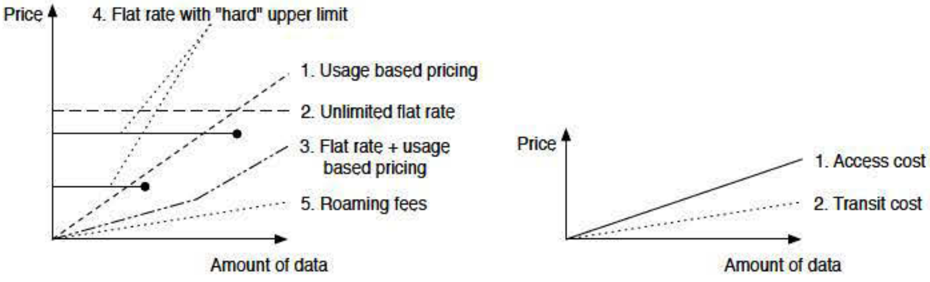
where D_i is the amount of data used during the time interval T_i, P_i is the price per data unit paid by the user for the data D_i and C_i is the production cost per data unit for the operator to deliver the data D_i.

4.3 Cost revenue function

The cost revenue function representing the usage based pricing showed in Equation (4.3) that was used in the simulation experiments is only one form for the cost revenue function. Different pricing models have different cost revenue functions. The unified (generic) form of the cost revenue function (f_{RC}) can be expressed as:

$$f_{RC} = f_R - f_C. \tag{4.4}$$

where f_R is the revenue function and respectively f_C is the cost function. For the revenue function, different pricing models including roaming fees when serving visited users needs to be considered. Therefore, when calculating the revenue function, it is also important to consider whose network resources are used and whether the served users are the operator's own users or visiting users. When using an operator's own network resources, then the operator also takes care of all production costs such as access and transit costs. Respectively, when serving visiting users, the operator receives roaming fees from the users' subscription providers.



(a) Revenue function examples. (b) Cost function examples.

Figure 4.1: Cost revenue function examples

There are different kinds of end-user pricing models that mobile operators use and four different revenue functions corresponding to different pricing models and roaming fees are shown in Figure 4.1(a). The roaming fees are not dependent on the pricing model, but they represent a source of revenue for an operator.

The first function corresponds to the “traditional” usage (or volume) based pricing where each data unit has a fixed price and no matter how much data has been transferred, the data unit price does not change. This is a quite straightforward and safe pricing model for operators, since all the transferred traffic generates profit. There is no issue of waste generation like with “basic” flat rate pricing [22]. For the end users, this model may not be the most predictable pricing model, since especially when we consider packet data services like the General Packet Radio Service (GPRS) or the Universal Mobile Telecommunications System (UMTS), a user has a hard time working out how much data has been transferred or how much data should be transferred to view, for example, a web page.

The second one is an unlimited flat rate where a user can transfer as much data as (s)he wants with a fixed fee. This pricing model is problematic for operators due to waste generation (Edell & Varayia [22]) where after a certain point, the transferred data represents a lower value for a user but still costs the same to an operator. Therefore, “enhanced” flat rate pricing models have been taken into use, like the third one, where the flat rate pricing ends when its transfer budget has been used and after that, the usage based pricing is used. This model ensures that all transferred data generates some profit and that no waste is generated. However the point where the flat rate limit has been exceeded and the usage based pricing is enabled is not so predictable for the end users.

The fourth example is a flat rate with an upper limit where a user pays for a certain data allowance, like 2GB [45]. This is very user-friendly model up to the limit and after this if the degradation is used, it does not yield any extra costs for the end users, but can result in drastic changes in data transfer speeds. The fifth example represents roaming fees that are already widely used by operators and represent traditional costs associated with mobile users roaming [29] [30]. There are different practices on how roaming tariffs are carried out like per-minute charging, data volume based charging and so on, and in our model, all monetary compensation (tariffs) coming from a home operator due to roaming are referred to as roaming fee.

The revenue function is naturally case sensitive, since different pricing models can result in different amounts of revenue for an operator under the same conditions and this is also visible in Figure 4.1(a). There might be a general concern that cooperation might result in increased prices due to the lack of competition. However, the regulation takes care of this and ensures fair competition in the market and that the Ambient Networks concepts enable users to freely select what networks they are going to use and this should prevent unfair pricing in a competitive market situation. When calculating the cost revenue function, it should also be remembered that an operator might use many different pricing models. For instance, packet data services could use some form of flat rate pricing whereas value added services could use another kind of pricing. So the revenue function f_R can also be the sum of multiple pricing model dependent functions and can be presented as:

$$f_R = EndUserFees + RoamingFees, \tag{4.5}$$

and

$$f_R = \sum_{i=1}^{K}(\sum_{j=1}^{L} f_{Rij}(data_{ij}) + \sum_{l=1}^{M} f_{Ril}(data_{il})), \tag{4.6}$$

where K is a number of the users, L is a number of the used pricing model dependent on revenue functions and M is a number of the used roaming fees. For instance, there could be a different roaming fee for voice services and data packet services. When calculating the costs of an operator's own users, then naturally the last factor in the formula is equal to zero, since there are no roaming fees involved.

Unlike the revenue function, the cost function is not pricing model sensitive, but there could still be different data transfer costs associated with different access networks. Also, for some traffic, an additional cost could be produced in the form of transit fees. For instance, costs could be different if the operator provides its own services or if the operator acts as a bit pipe and only transfer user traffic toward the Internet via a transit service provider. Figure 4.1(b) illustrates a simple view of different cost factors. Access costs are the operator's internal costs depending on the operator's network deployment and configuration. Transit costs are from the Internet and the Internet Service Provider (ISP) "world" representing for example transit service costs between two autonomous systems. It should be noted that not all transit traffic generates transit costs. Transit costs in Internet and related business models are further described in [56].

So according to this view, the generic form of the cost function can be expressed as:

$$f_C = AccessCosts + TransitCosts, \tag{4.7}$$

and

$$f_C = \sum_{i=1}^{K} (\sum_{j=1}^{L} f_{Cij}(data_{ij}) + \sum_{k=1}^{M} f_{Cik}(data_{ik})), \tag{4.8}$$

where K is a number of the users, L is a number of the used accesses and M is a number of the used transit providers.

Chapter 5

Simulation model

The developed access selection algorithm also needs to be assessed and this is why the simulation model is defined. The model is used to simulate both a single operator multi-access mobile network as well as a multi-operator environment. The simulation model covers the needs of both "basic" and "advanced" algorithms. The former does not exploit any handover model and application session and radio interface bootstrapping delays are omitted, thus they are only used by the "advanced" model. So the same model is used with different configurations for the algorithm evaluations. Both simulation experiments use the same approach: i) perform a simulation run without the network cooperation and then ii) repeat the simulation with the network cooperation.

This chapter presents the simulation model used for the algorithm evaluations. For both "basic" and "advanced" algorithms, there are some common working assumptions that are explained in Section 5.1. These two algorithms use different sets of the decision making constraints and these sets are explained in this section. The single operator specific settings are then explained in Section 5.2 and correspondingly Section 5.3 describes the multi-operator specific settings.

5.1 Working assumptions

For all simulations, the simulation area is 1 square kilometre and the deployment of the radio access networks is simulation case specific, i.e., the single operator simulations use different deployment than the multi-operator ones in terms of a number of used Radio Access Technologies (RATs). The radio cell deployment is partially randomized based on the uniform distribution. The simulation area is divided into 16 regions that are each 250m * 250m.

These regions are used to help the cell deployment phase. For each region, a number of radio cells are defined and for some regions this number could be also equal to zero. Once this is done, for each region hosting radio cells, the cell origin locations are randomized within the region based on the uniform distribution. After this, cell radius is defined based on the pre-configured info, i.e., all radio cells of the same type use the same cell radius. There is one exception, the single operator simulations use different Wireless LAN (WLAN) cell sizes.

The access network and cell load capacity is expressed by using Traffic Unit (TU), which is an abstract measure to express traffic load; how much traffic is generated by a mobile node. For each mobile node, there are two sets of vectors defined: i) movement patters and ii) application usage patterns. All vectors are randomly generated based on the uniform distribution. For the movement vectors, the following working assumptions are used:

- Starting locations are randomized based on the uniform distribution,
- The value range of speed is $]0, 10]m/s$, and
- A random ±90° movement direction change probability is used.

The application usage patterns define what application type a mobile node tries to use and when. There are no idle moments for the application sessions, thus all mobile nodes try to use an application all the time. Whether a mobile node is able to use the requested application type, depends on the mobile node's location and available access resources.

Other working assumptions are as follows:

- Each mobile node supports all used RAT types, e.g., WLAN and Universal Mobile Telecommunications System (UMTS).
- Only one radio access could be in use at a time.
- Each mobile node supports all used handover types including handovers between access networks as well within access networks.
- Cell loads are measured using TUs generated by associated mobile nodes.
- Cells' coverage areas are circular and signal strength S is defined as

$$S = \max[0, 1 - (d/R)], \tag{5.1}$$

where d is the distance between a mobile node and the cell origin and R is the cell radius.

5.2 Single operator environment

For a single operator setup, three RATs are used; 1 Wireless Wide Area Network (WWAN) like UMTS and two short range RATs like WLAN. Table 5.1 shows a number of different types of radio cells including their capacities (in term of TUs) and the cell radius and respectively radio cell deployment for the single operator experiment is presented in Figure 5.1. It should be noted that WLAN cell sizes were not fixed and the maximum radius sizes are listed in the table.

RAT type	number of cells	cell capacity	cell radius
WWAN	4	30TUs	300m
WLANa	30	5TUs	up to 31m
WLANb	60	5TUs	up to 72m

Table 5.1: RAT specific configurations

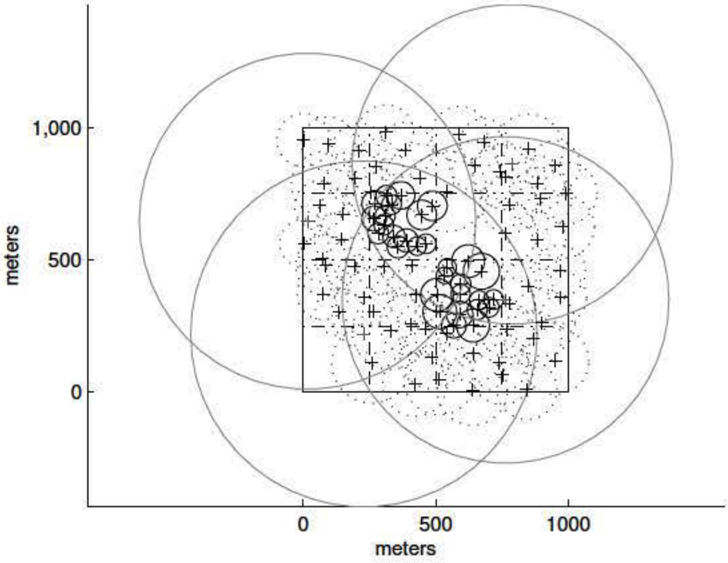


Figure 5.1: Network topology

There are two types of traffic, the first type (best effort) requires 1 TU of the access resources and the second (best effort plus) needs 50% more, i.e., 1.5 TUs. A preferred traffic type is defined by application vectors that are calculated for each mobile node and that represent what type of service a mobile node requests each time. Based on the traffic type, a specific set of constraints is used meaning that the handled traffic type affects how accesses are evaluated and selected.

There are no explicit rules for how to select what decision making constraints are used. Basically it is question of what kinds of aspects the access selection decision making should be able to differentiate between and also the simulation model used has an impact, i.e., what kinds of functionalities it supports. Signal strength is a natural choice, since it defines what can be detected. Cell load levels are included because one of the goals was to include a cell level load balancing. These form a kind of minimal set that is used for best effort traffic type. For best effort “plus”, some additional constraints are selected. The selection of RAT is used to control inter-RAT handovers and two traffic type constraints are used to take service type related aspects into account. These constraints are sufficient to study the access selection decision making, which is RAT and traffic or service type sensitive and also supports load balancing.

For the best effort type of traffic Table 5.2 shows the constraints used in the decision making, their types and their constraint weights. The first constraint is the terminal constraint and another is the network constraint. The terminal constraints are handled in a terminal and correspondingly the network constraints in the network side.

Constraint name	Constraint type	Weight
Signal strength	non-binary/soft/terminal	1.0
Cell load levels	non-binary/soft/network	1.0

Table 5.2: Constraint types for best effort traffic

Respectively, for best effort plus type of traffic Table 5.3 shows the used constraints in the decision making, their types and their constraint weights. The three first constraints are the terminal constraints and the rest are the network constraints.

For these single operator simulation experiments, only soft constraints are used. Additionally, *Cell load levels* constraint is used even under a light network load enabling the load balancing to be done according to the cell load levels.

Signal strength constraint prefers a stronger radio signal. This is perhaps one of the most significant constraints used in legacy systems to perform access selection. In our model, it is used just as any other constraint without any special handling, apart from having a relatively high weight.

Constraint name	Constraint type	Weight
Signal strength	non-binary/soft/terminal	0.4
Selection of RAT	binary/soft/terminal	0.3
Used traffic type	binary/soft/terminal	0.3
Cell load levels	non-binary/soft/network	0.7
Supported traffic type	binary/soft/network	0.3

Table 5.3: Constraint types for best effort "plus" traffic

Selection of RAT constraint prefers the discovered cells that are in the current RAT and it is used to minimize inter-RAT handovers.

Used traffic type constraint prefers the use of discovered cells that supports best effort plus traffic type.

Cell load levels constraint is used for load balancing and this constraint prioritizes the cells with lower load over the highly loaded ones.

Supported traffic type constraint is used to prioritize detected radio cells according to their supported traffic type compared to the user's requested traffic type.

These constraints are used for the new access selection strategies and the legacy strategy uses only *Signal strength* constraint.

5.3 Multi-operator environment

In the multi-operator setup, two operators and two service providers are used. Both operators had a Service Level Agreement (SLA) with their service provider and they provide the same RATs, one access network with 45 WLAN cells and another with 2 UMTS cells. The WLAN cells have a radius of 80m and UMTS cells have a radius of 600m. The simulation area is one square kilometre. Radio cell deployment in the simulation area is presented in Figure 5.2 where one operator's cells are presented with grey dashed lines and another one's with black solid lines. The number of mobile nodes varies between 100 and 1000 mobile nodes depending on the simulation measurement used.

If a mobile node requests a service type that is not available at the mobile node's current location, then the mobile node goes to the *disconnected* state and releases any reserved access resources. A mobile node stays in the *disconnected* state until sufficient access resources and requested service types are available.

Figure 5.3(a) illustrates the multi-operator setup when the cooperation

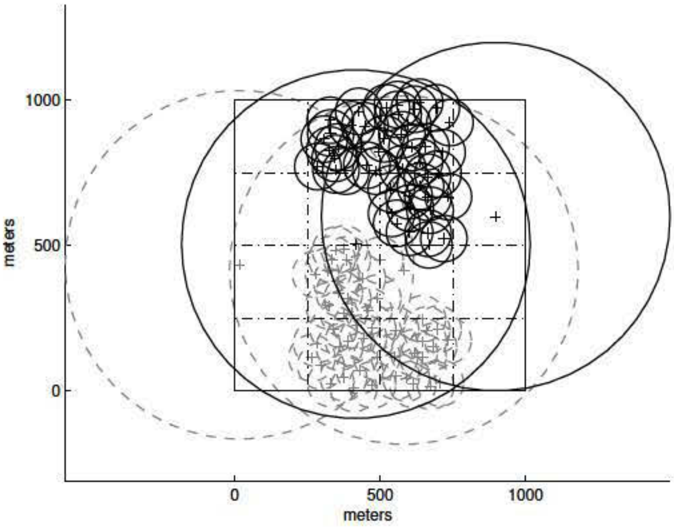
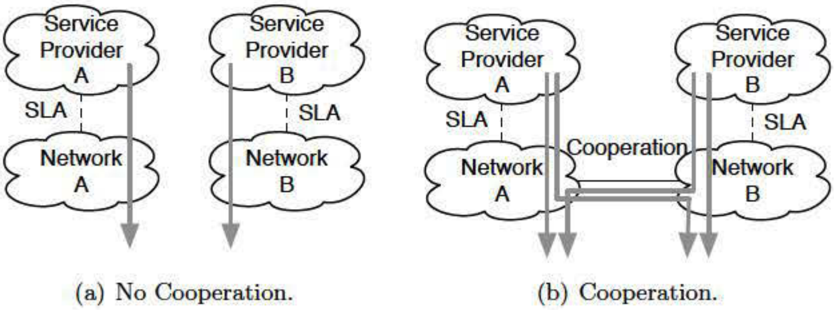


Figure 5.2: Network topology



(a) No Cooperation.

(b) Cooperation.

Figure 5.3: Setup with and without the cooperation

is not supported. In this case, both operators are limited to offering access services only for their own subscribers and the services are from the service provider with whom they have an SLA. When the network cooperation is present, the situation is different as shown in Figure 5.3(b), where the cooperation enables three additional features; i) serving visited users, ii) providing external services and iii) supporting handovers between operators.

The handover model combines the radio and application connectivity states. The model consists of five states as shown in Figure 5.4; *disconnected*, *connected*, *session association*, *radio bootstrapping* and *handover* *execution.*

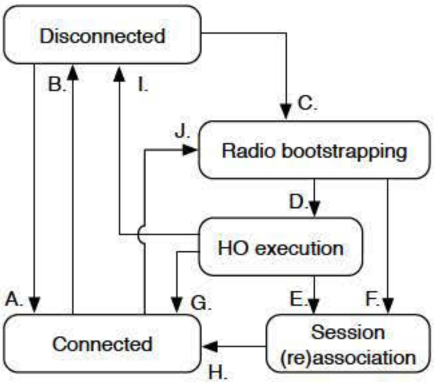


Figure 5.4: State machine

Table 5.4 explains under which conditions the state transitions occur in this model. All applications use the same session association delay of one time unit. For UMTS, it was assumed that this radio technology is attached all the time due to its low power consumption compared to WLAN, which is kept de-attached, if not in use. Therefore, the bootstrapping of UMTS does not incur any additional delay, thus its radio association is considered to be an instant action from the simulation perspective. The bootstrapping time of WLAN was set to one time unit.

Transition	Conditions
A	Simulation start-up.
B	Out of coverage or no available access resources.
C	New radio access discovered.
D	New radio access ready.
E	Handover finished successfully and a session needs to be (re)associated due to the handover type or due to the application type change during the handover.
F	The discovered new radio bootstrapped successfully.
G	Handover finished successfully and no need to re-associate a session.
H	Session (re)associated successfully.
I	Handover initiated with radio bootstrapping.
J	Handover failed and the old cell not available.

Table 5.4: State transition conditions

For each type of handover, a handover execution delay is defined as illustrated in Table 5.5. Using the UMTS and WLAN performance results from [48] as references, a “basic” handover type within a single RAT, intra-RAT handover, is chosen to last for one time unit. Other types (inter-RAT and inter-operator handover) are chosen to last twice as long as the delay of the intra-RAT handovers. Intra-RAT means a handover inside the same RAT and respectively inter-RAT means a handover between RATs of the same or different types. For instance, inter-RAT WLAN handovers could happen between two WLANs for the same operator, like between WLANa and WLANb that are listed in Table 5.1. It should be noted, that not all presented handover types are supported in all simulation cases. For instance, in the multi-operator environment (Section 5.3) inter-RAT WLAN and inter-RAT UMTS types are not possible, since operators have not multiple RATs of the same type.

Handover type	Delay (time units)
Intra-RAT WLAN	1
Intra-RAT UMTS	1
Inter-RAT WLAN-UMTS	2
Inter-RAT UMTS-WLAN	2
Inter-RAT WLAN	1
Inter-RAT UMTS	1
Inter-operator WLAN-UMTS	2
Inter-operator UMTS-WLAN	2
Inter-operator WLAN	2
Inter-operator UMTS	2

Table 5.5: RAT specific handover execution delays

There is one exception to how these delays are applied; when a simulation is started, the bootstrapping delays are not used, thus all mobile nodes are able to move directly into the *connected* state assuming that enough radio access resources and a requested service type are available. So from an end-user perspective, the overall effective handover execution time is the sum of a handover delay according to Table 5.5, a radio bootstrapping time and a session association time.

In the simulation model, inter-RAT handovers are not supported by the legacy access selection strategy, which is always forced to perform a radio re-association when switching between RATs. In this way, the legacy strategy provides a good reference point by which the two new access selection strategies can be assessed. In the used simulation model, this means that a mobile node is disconnected for a short period of time. The network-centric and terminal-centric strategies support Inter-RAT handovers inside an operator network and also between operators' networks if the cooperation exists.

Compared to the single operator case (Section 5.2), some additional constraints are needed due to multi-operator environment. In order to control handovers between operators, a selection of operator constraint is included. Respectively, a roaming agreement constraint is needed to differentiate between a visited and home networks. In addition, some constraints are set to be hard constraints now compared to the single operator case where all constraints were soft ones. For the roaming agreement constraint, this means that a user cannot connect to a visited network unless it has a roaming agreement with the user's home operator. For the supported service type constraint, this mean that unless the requested service type is not supported, a device cannot connect to a radio cell. These changes were necessary because it is a multi-operator environment and this simulation case is also used for the calculation of the new performance metrics involving a collection of the satisfaction related data.

Table 5.6 lists the used constraints, their types and their constraint weights. The three first constraints are the terminal constraints and the rest are the network constraints. The terminal constraints are handled in a terminal and correspondingly the network constraints in the network side.

Constraint name	Constraint type	Weight
Signal strength	non-binary/soft/terminal	0.6
Selection of RAT	binary/soft/terminal	0.3
Selection of operator	binary/soft/terminal	0.1
Cell load levels	non-binary/soft/network	0.6
Roaming agreement	binary/hard/network	1
Supported service type	binary/hard/network	1
Service load levels	non-binary/soft/network	0.4

Table 5.6: Constraint types

As illustrated in Table 5.6, the constraint specific weights are only defined for the soft constraints. The hard constraints that are used for qualifying evaluated cells according to the constraint's conditions, the weight value 1 is used in the cell rank calculation. The sum of all terminal or network soft constraints is equal to 1. It should be noted that to avoid an unnecessary *ping pong* effect between the cells with a light load, a load balancing threshold of 80% was used. In other words, all cells with less than 80% load are considered to be equally loaded.

Signal strength constraint prefers a stronger radio signal. This is perhaps one of the most significant constraints used in legacy systems to perform access selection. In our model, it was used just as any other constraint without any special handling, if only it had a relatively high weight.

Selection of RAT constraint prefers the discovered cells that are in the current RAT and it is used to minimize inter-RAT handovers. Respectively *Selection of operator* constraint prefers the discovered cells from the current operator. In non-cooperative case, naturally this constraint has no role, since mobile nodes are not able to use more than one operator's access networks.

Cell load levels and *Service load levels* constraints are used for load balancing. The former is used to prioritize the cells with lower load over the highly loaded ones assuming that cells' load levels exceed the load balancing threshold. The latter does the same for service types.

Roaming agreement and *Supported service type* are both binary hard constraints and they are used to disqualify the cells that belong to the operator either not having a valid Roaming Agreement (RA) or not supporting the requested service type.

These constraints are used for the new access selection strategies and the legacy strategy uses only *Signal strength* and *Roaming agreement* constraints. The constraint specific weights are used to define how significant constraints are compared to each other. For instance, if an operator wants to fully support load balancing, then cell load level constraint's weight should be increased while other network constraints weight are decreased. In the same way, if an operator would like to use access selection decision making to help RAT level load balancing, a new constraint could be defined which would represent the load level of RAT.

5.4 Experiment

For a multi-operator (Section 5.3) environment, the simulation experiment follows the two-step approach, where i) a simulation experiment without the network cooperation is first performed and then ii) a configuration is changed so that the network cooperation is enabled and the simulation is repeated. This results in two sets of measurements for all simulated access selection strategies, one with the network cooperation and another without. These measurements are comparable with each other. In the case of single operator (Section 5.2) simulations where the network cooperation is not included, the latter step is not needed. A single simulation experiment contains 1200 time units where 1 seconds corresponds to 10 time units. So one simulation experiment lasts 120 seconds and measurements are done 10 times per second, i.e., 1 measurement in each time unit.

For the single operator simulations, where a number of users is scaled from 100 to 1000 users, the experiment is repeated for each measured number of users; $[100, 200, \ldots, 1000]$. The movement and usage patterns are generated only once meaning that these patterns are used for each measured cases. The number of users is scaled to study how different access selection strategies perform under different network loads to show their scalability in terms of network load. The same scalability study is done for the multi-operator simulations as well, but the upper limit of users is 800 users in this case. In both cases, single operator and multi-operator, 300 users are used as a reference case where other kinds of measurements related to users (dis)connectivity and handover numbers is done. 300 users are selected since it represents a lightly overloaded network situation, where there is still enough available network resources for a load balancing.

For each described configurations, the simulation experiments were done once. The same model and implementation were used in other experiments where the access selection performance was studied from different perspectives compared to this research [48] [61] [67] [81]. In those experiments, different sets of the decision making constraints were used. This naturally means that their results are not directly comparable with the results presented in this dissertation. However, the findings and the results are in line with the results of the presented simulation experiments and confirms the performance differences between the studied strategies.

The results are expected to show performance differences between the studied access selection strategies. It was also interesting to see what would be the role of network cooperation; would it have any kind of visible impact? The results should help to answer the described research questions:

- What kinds of technical benefits could a new distributed access selection algorithm exploiting network cooperation provide?
- How to model and measure the end-user satisfaction and how to measure aspects related to the operator satisfaction.
- How did the gained technical benefits translate into enhanced end-user satisfaction?

Chapter 6

Analysis of the simulation results

This chapter presents the results and analysis of the simulation experiments. The results show that for the service availability in a single operator environment, the average online time is about 10% higher when the new access selection strategies are used compared to the legacy one. When the allowed offline time percentage for the active mobile nodes is required to be low, i.e., less than 5%, then the new access selection strategies provide substantial gains over the legacy one. For the network utilization in a multi-operator environment, the network-centric strategy results in about a 20% higher utilization rate over the legacy one when the network load is not too high. Both in single and multi-operator cases, a capability for supporting off-loading traffic from cellular networks to Wireless LAN (WLAN) hot spots has an important role and is one of the main reasons why the new access selection strategies outperforms the legacy one.

The two new access selection strategies are also able to benefit better from network cooperation as measured by the users' disconnectivity. For the cost revenue function, the results show that the network cooperation results in a 20% higher value. The single operator multi-access simulation results for the access selection strategies using the "basic" algorithm version are shown in Section 6.1. Respectively, the results for the "advanced" algorithm in a multi-operator multi-access environment are presented in Section 6.2. Section 6.3 shows the satisfaction measurements for the multi-operator multi-access simulations using the Operator Satisfaction Indicator (OSI) and User Satisfaction Index (USI) calculations. Finally, the main findings are summarized in Section 6.5.

6.1 Single operator results

The main goal of these single operator simulations is to study what kinds of (technical) benefits the new distributed access selection strategies defined in Section 2.4.3 can potentially bring. The network cooperation aspects are not measured explicitly, thus sufficient cooperation is present for the two new access selection strategies that are able to exploit it. Additionally, the theoretical upper boundaries are also provided in order to assess how well the new strategies perform.

Figure 6.1 illustrates the average connectivity time for a single user for all three evaluated access selection strategies under different network loads when the number of mobile nodes varies between 100 and 1000. The figure also includes the optimal graph showing the theoretical upper boundary. The optimum solution for each number of mobile nodes was calculated separately using the same MATLAB®simulation setup in order to produce comparable results. For the optimum solution calculations, the Linear Programming (LP) technique called Mixed Integer Programming (MIP) [39] was used[1]. As the results clearly show, as the number of mobile nodes increase, the network becomes overloaded, which was expected, and for 1000 mobile nodes, a user spends more than half of the simulation time disconnected due to a lack of access resources.

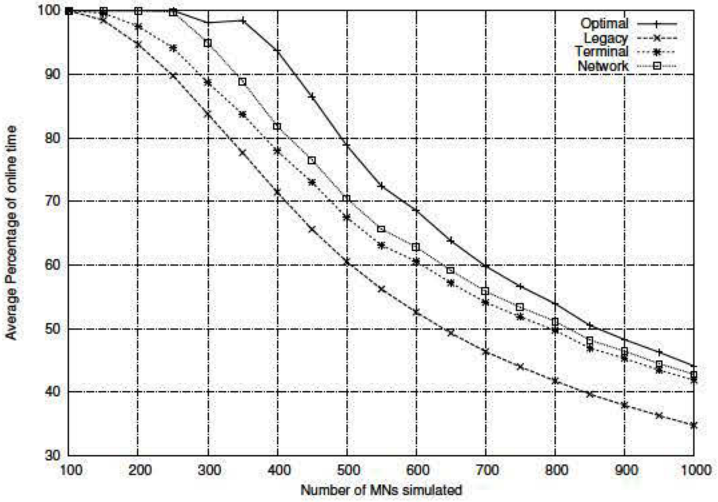


Figure 6.1: Average connectivity time [PI]

---

[1] In Figure 6.1 and Figure 6.2 there are some abnormalities in the optimum curve that most likely are a result of programming error(s) or other error(s) in the calculations.

The network-centric strategy outperforms the other two under all network loads. However as the optimal curve shows, there is still room for improvement and this delta varies between 10% (with low network loads) and 2% (with high network loads). One important finding is that the difference between the legacy and the network-centric case is fairly constant. It is also interesting how the terminal-centric strategy behaves; with a light load its performance is close to the legacy strategy but once the network load increases it starts to increase the gap to the legacy strategy and to approach the performance of the network-centric strategy. The reason for this is that the network-centric strategy focuses more on the load balancing of the access resources than the terminal-centric strategy, i.e., it is off-loading traffic from cellular to WLAN when feasible and this frees up resources for mobile nodes that are not in the coverage of WLAN and with a light network load there are enough available resources for the efficient load balancing, which is not the case with a high network load.

These connectivity measurements could also be seen from a different perspective as shown in Figure 6.2. The figure shows how different access selection strategies perform in terms of the allowed offline time of an average user. Once again, the network-centric strategy clearly outperforms the other two with a low allowance by being able to serve more mobile nodes, this can also be seen as an effective use of network capacity. This difference gets smaller between the network-centric and terminal-centric strategies when the allowance increases, but is remaining fairly constant compared to the legacy case. This is well in-line with the results shown in Figure 6.1.

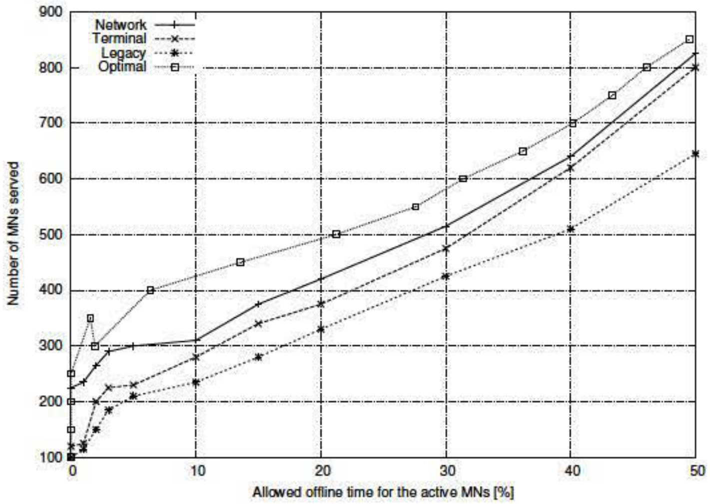


Figure 6.2: Average offline time allowance [PI]

The improved effective network capacity that resulted in the use of the network-centric access selection strategy does not come without a high price as shown in Figure 6.3, where an average number of performed handovers per mobile node under different network loads are presented. The network-centric strategy uses three times more handovers than the legacy case with 300 mobile nodes. However this number becomes quickly lower as the network load increases and with 500 mobile nodes, the number of handovers has already more than halved. This is the result of a "ping pong" effect that occurs in lightly loaded access networks, i.e., a handover is initialized even when there is no imminent lack of resources in the old radio cell. This is fixed in the "advanced" algorithm that is used for the multi-operator simulations by introducing a load balancing threshold parameter for the load level and load balancing is enabled when this level is reached. In practice, this means that for lightly loaded cells, the load balancing is not factored in the decision making. However in order to do load balancing at the radio cell level, handovers are mandatory tools for the load balancing, thus an effective load balancing results in additional handovers.

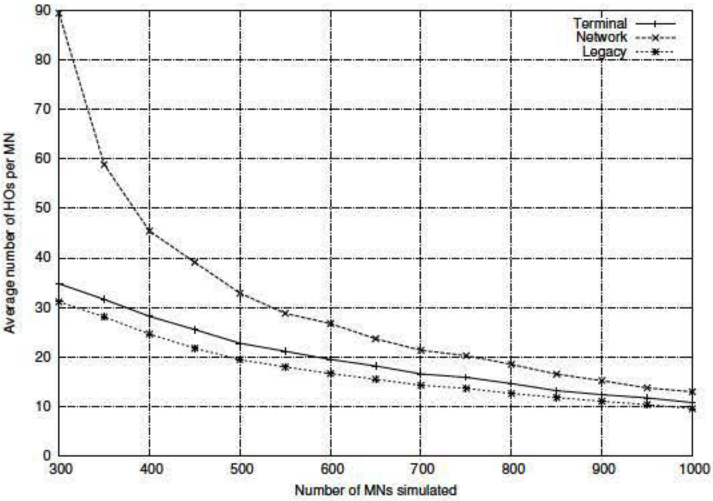


Figure 6.3: Average handovers per mobile node [PI]

Respectively, Figure 6.4 further details the handover statistics by showing different types of handovers for 300 mobile nodes. The network-centric strategy uses twice as many intra-Radio Access Technology (RAT) handovers (horizontal handovers) than the two others. The difference is even bigger when inspecting inter-RAT handovers (vertical handovers). However, it is important to see that for the terminal-centric and legacy cases, almost the same number of different types of handovers is used. Because this indicates that it is still possible to improve the access selection performance without substantially increasing a number of required handovers.

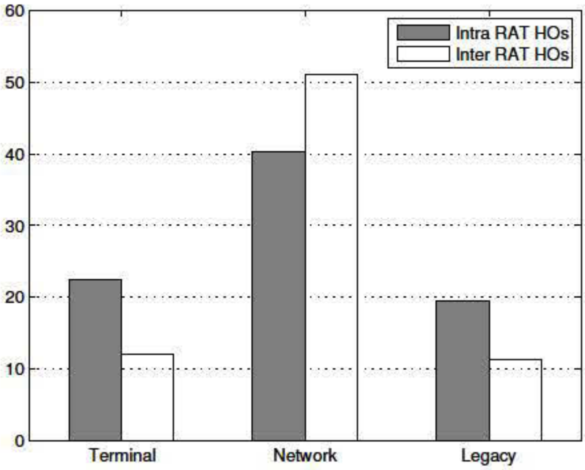


Figure 6.4: Average handovers of different types per mobile node (300 mobile nodes) [PII]

A relatively high number of vertical handovers that was required by the network-centric strategy is not a good sign. Because typically (seamless) vertical handovers where users do not experience any disruptions during handovers are technically more challenging to implement than horizontal ones [79]. This also implies that a cost of a vertical handover is higher than a cost for a horizontal handover in terms of required signalling and processing resources. How much higher this cost is depends on various things like what type of mobile network architecture we are considering and what kind of network configuration is in use. For instance, if a vertical handover is done from one access network to another, then this could result in state update(s) in core network side. In such case, naturally a higher number of network elements could be involved than in a simple horizontal handover where a mobile device just switch from one cell to another. From a mobile network point of view, this means more control signalling for which more processing capacity is required. Vertical handovers can also increase handover delays [19] [20] because when for instance Mobile IP is used then a handover event is not anymore transparent to the network layer (IP) and therefore there are some extra things to be done in a mobile device like IP address configuration, routing table updates and binding updates. Naturally in this kind of cases vertical handovers would not be so favorable thing to do, thus their usage should be minimized. Longer handover delays would naturally hinder the performance of the network-centric strategy, since longer delays would most likely affect average Quality of Experience (QoE), especially if we consider quality sensitive traffic.

All this increased signalling and processing load to handle additional handovers could translate into an increased Operating Expenditure (OPEX) for an operator. From a user perspective, an increased number of vertical handovers could mean a shorter battery lifetime due to an increased power consumption when a mobile device needs to keep multiple radios enabled. Also short delays in data transfer can be possible when context switching is done between different radio bearers. Whether this increased number of vertical handovers is acceptable, it depends on how effectively these vertical handovers can be done and what kind of extra investments this would require from an operator. For users, an added cost is perhaps easily acceptable, since that is a price for being able to be "always best connected". On the other hand, if there are additional expenses (Capital Expenditure (CAPEX) or OPEX) for an operator, this might become visible for end users if pricing is changed accordingly, for instance by increasing voice or data rates.

The single operator results show clearly how the two new access selection strategies perform better than legacy one. New strategies scale better than the legacy one, but this does not come without additional cost, which is a higher number of needed handovers, especially in the network-centric strategy. Handovers are needed to (re)arrange users so that the available resources can be utilized effectively. Whether this additional cost is significant or not, depends on how costly we consider handovers to be. If the cost of a single handover is taken to be low, then this increased need for handovers is not so significant.

6.2 Multi-operator results

Compared to the single operator measurements, the simulation scope is now wider, i.e., network cooperation is included in it. The new distributed access selection algorithm (Section 2.4.4) should be better at exploiting available network cooperation compared to the legacy one. This ability could increase the effective usage of network resources. These simulations study how visible this ability is in the results and how well the legacy strategy handles the network cooperation. So from the access selection perspective, it is important to know; "Does it matter whether network cooperation is present or not and if it is how much better is it?"

First, in order to study how the used access selection strategy affects network scalability with and without the network cooperation, the best and worst access selection strategy based on the other measurements are compared. Figure 6.5 shows how the network-centric and the legacy strategies are able to serve users under increasing network loads. For 100 mobile nodes, it does not matter whether network cooperation is present or not. Its significance becomes more clear when the network load increases. When the network load is between 200 and 800 mobile nodes, the difference with these network loads is approximately constant for both evaluated cases. Also the legacy strategy benefits having network cooperation, but the improvements gained by having it are not as big as for the network-centric strategy. This is also visible when comparing strategies with and without the network cooperation, i.e., the difference is bigger with cooperation than without it. As the results clearly show, the network-centric strategy maintains better its capability to serve users under a heavy network load.

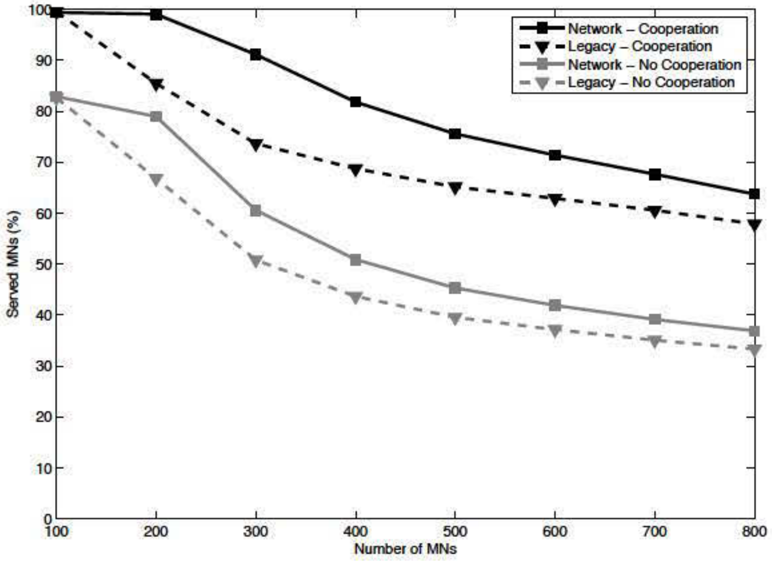
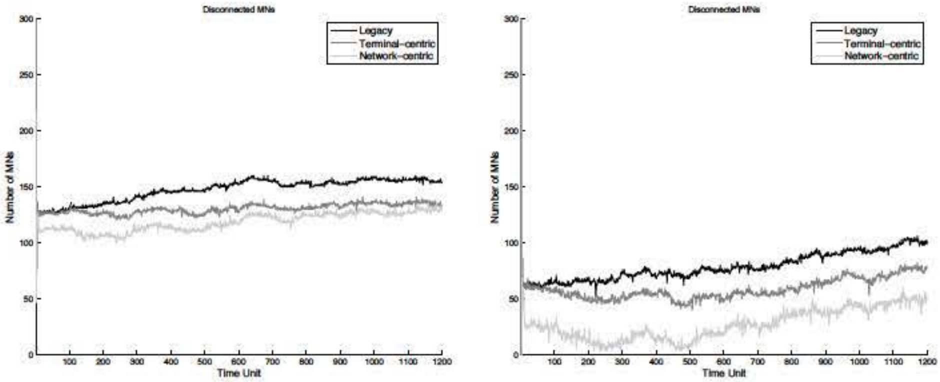


Figure 6.5: Served mobile nodes [PIII] [PV] [PVI] [PVII] [PVIII]

Even though the differences between the cases with and without the cooperation get smaller under a heavy load, the network-centric strategy still yields approximately 20% improvement in network utilization (in terms of the number of served users) compared to the legacy strategy with 300 users when the cooperation is supported. When the cooperation is missing, then the network utilization increases slightly less, which is due to a lack of extended access coverage and supported services. For 300 users, the network-centric strategy results in approximately 30% higher utilization with the cooperation than without it. Also these results show how important role traffic off-loading plays and they confirm the findings of the single operator simulation experiments and the behaviour of both evaluated cases is very similar compared to the single operator results.

For 300 mobile nodes, Figure 6.6(a) shows the number of disconnected users in the simulation for each evaluated access selection strategy without the cooperation. The same trend that is visible in the single operator results is also visible here, both the network-centric and the terminal-centric strategies perform better than the legacy one. For instance, the behaviour of the terminal-centric strategy is very similar than in Figure 6.1 and in Figure 6.2, i.e., its performance is approaching to the network-centric case towards to the end of the simulation.



(a) Disconnectivity - no cooperation. (b) Disconnectivity - cooperation.

Figure 6.6: Disconnectivity for 300 mobile nodes [PVIII]

A high number of disconnected mobile nodes is the result of two factors. Firstly, the lack of the network cooperation limits the available operators and services, i.e., there is only one operator with one service type. Secondly, since the WLAN hot spots are not uniformly distributed as explained in Chapter 5, the load balancing and traffic off-loading cannot be done effectively everywhere leading to temporary congestion in some areas, i.e., only approximately 38% of the simulation area is populated by WLAN hot spots.

The corresponding results with the network cooperation between operators are shown in Figure 6.6(b). All three access selection strategies benefit from the presence of the cooperation, which was expected. The network-centric and the terminal-centric strategies can better exploit the network cooperation as shown in the figure. This is the reason why for instance the gap between the network-centric and the legacy cases is bigger than it was without the network cooperation.

As in the single operator case, an improved performance gained by the use of new access selection strategy does not come without a price, i.e., an increased number of handovers is needed. The handover measurements are presented in Table 6.1 and they follow a similar trend as was observable in the single operator results, i.e., a better performance involves a higher number of handovers. Once again, the network-centric strategy results in the highest number of any kind of handover. Without the cooperation, the network-centric case required about 2.3 times more intra-RAT handovers than the legacy one and for the terminal-centric case, this number was 1.7. Correspondingly, with the cooperation these numbers were 2.2 for the network-centric case and 1.5 for the terminal-centric case. For inter-RAT handovers, the differences between the network-centric and the terminal-centric cases are not so significant, but what is noteworthy is that the network-centric case required almost twice as much inter-operator handovers than the terminal-centric case.

		Intra-RAT	Inter-RAT	Inter-oper.
No Coop				
	Legacy	1.63	0	0
	Terminal	2.77	0.29	0
	Network	3.76	0.36	0
Coop				
	Legacy	4.45	0	0
	Terminal	6.60	0.51	0.41
	Network	9.90	0.56	0.77

Table 6.1: Handover statistics (avg. handovers per mobile node) [PVIII]

The most significant difference in these results compared to the single operator results is the performance of the terminal-centric strategy, which is now clearly worse compared to the legacy case in terms of the number of required handovers. These results indicate that additional handovers are indeed needed to support efficient load balancing and traffic off-loading.

Whether the gained benefits are big enough to justify increased signalling and processing requirements, it depends on what kind of cost is associated with different types of handovers. The table also shows how the legacy strategy does only support handovers within a single RAT, but this was one of the design choices for the legacy strategy as explained in Section 2.4.

6.3 User satisfaction results

The satisfaction measurement consists of 3 different cases, where each case uses its own mobile node movement patterns. Two of these cases use 300 mobile nodes and one uses 200 mobile nodes. Therefore, there are some minor differences in the numerical results, but all three results are consistent and support each other.

Table 6.2 shows the P class distribution for 300 mobile nodes based on which the USI calculations in Table 6.3 using 3 different P class weight sets were done. When the network cooperation is present, there are over a 30% higher number of connected users (183 connected users vs. 273 connected users). Without the cooperation, users spend almost 40% of their time disconnected and with the cooperation the corresponding time is around 9%.

	Not Connected	Connected		
	P1	P2	P3	P4
No Coop	39.4%	23.6%	14.6%	22.5%
Coop	8.8%	5.7%	25.3%	60.3%

Table 6.2: Network-centric - P class distribution [PIV] [PVI]

Depending on the used P class weight set, this translates into different amounts of increased user satisfaction as illustrated in Table 6.3.

P class weights $[\alpha, \beta, \chi, \delta]$	No Coop	Coop	Ratio
$[0, 1, 1, 1]$	726.66	902.58	1.2
$[-1, 1, 1, 1]$	254.74	797.38	3.1
$[-1, 0.25, 1, 1.4]$	150.31	966.26	6.4

Table 6.3: Average USIs for different P class weight sets [PIV]

The first weight set ($[0, 1, 1, 1]$) does not include a negative weight value for disconnectivity and therefore the difference between having the cooperation and not having it is the smallest. The second set ($[-1, 1, 1, 1]$) uses a negative weight value for disconnectivity, i.e., for $P1$, but does not differentiate between $P2$, $P3$ and $P4$ classes and it results in lower USI results compared to the first set. The third set ($[-1, 0.25, 1, 1.4]$) differentiates between all P classes and results in both the lowest USI without the cooperation and the highest USI with the cooperation compared to other two. Also, the third set results in the highest USI value difference between "no cooperation" and "cooperation" cases. The relatively low USI for the third set when the cooperation is not present is the result of the high number of disconnections ($P1$) and the use of negative weight for these. Respectively, by using a higher weight value for "better than requested" ($P4$) measurements results in the highest calculated USI value. In practice, this means that the $P4$ weight value outweights the value of $P2$.

The "ratio" column shows the ratio between the USI value with and without the cooperation and the third set has over two times higher ratio compared to the second one. The first set clearly has the worst ratio and it resulted in a 1.2 times higher average USI per user with the cooperation than without it. For the second set, this ratio was 3.1 and respectively for the third set 6.4.

For each evaluated access selection strategy, the USI measurements are presented in Table 6.4. These calculations are based on the 300 mobile nodes set also used for the technical results in Section 6.2 and the second P class weight set ($[-1, 1, 1, 1]$) is used. The table also includes the number of connected mobile nodes and the normalized USI values.

	Connected Mobile Nodes (MNs)	USI	Norm. USI
Legacy - No Coop	51%	21	2
Terminal - No Coop	57%	158	13
Network - No Coop	61%	255	22
Legacy - Coop	73%	569	47
Terminal - Coop	81%	733	61
Network - Coop	91%	988	82

Table 6.4: Connection and USI statistics for connectivity sensitive users (300 mobile nodes) [PV] [PVIII]

The USI value normalization is done so that the maximum normalized value is 100 and it corresponds to the maximum absolute value, which is 1200 ($= 1200 * P4$ *weight*). Respectively the minimum normalized value -100 corresponds to the minimum absolute value -1200 ($= 1200 * P1$ *weight*).

An increased utilization rate and decreased mobile nodes disconnectivity showed in Section 6.2 naturally affect both access and service availability and these increase user happiness as it can be seen from the USI calculations. In practice, the second P class weight set represents user behaviour where the user is always equally happy whenever connected, i.e., different quality levels are not modelled. The negative weight value of $P1$ (disconnection) results in high USI value differences between different cases.

Respectively, Table 6.5 shows the USI measurements for the third P class weight set ($[-1, 0.25, 1, 1.4]$). The normalization is done in the same way with the exception that the value range of the absolute value is now [-1200,1650]. The negative weight value of $P1$ (disconnectivity) and a relatively low $P2$ weight value result in a negative USI value for the legacy case when the cooperation is not supported. This is fixed when the cooperation is added, i.e., a higher USI value; -126 vs. 740. As expected, both new access selection strategies outperform the legacy one, but their performance is also relatively poor without the cooperation. Compared to the Table 6.4 USI values, all normalized values are lower when the quality aspects are included. Now also the difference between the network-centric and terminal-centric is a bit smaller compared to the second P class weight set presented in Table 6.4.

	Connected MNs	USI	Norm. USI
Legacy - No Coop	51%	-126	-8
Terminal - No Coop	57%	37	2
Network - No Coop	61%	150	9
Legacy - Coop	73%	740	45
Terminal - Coop	81%	949	57
Network - Coop	91%	1227	74

Table 6.5: Connection and USI statistics for connectivity and quality sensitive users (300 mobile nodes) [PVIII]

For the network-centric and legacy strategies using the second P class weight set ($[-1, 1, 1, 1]$) and 200 mobile nodes the connectivity and normalized USI values are shown in Table 6.6. Comparing the presented USI calculations for 300 mobile nodes, there are no surprises and the results are well in-line with the other results. Perhaps the most significant difference is that now the differences between strategies with and without the cooperation are smaller. This is the result of having better connectivity due to a lighter network load and therefore even the legacy strategy is able to serve the users fairly well. The network-centric strategy achieves the normalized USI value 98 and the difference between strategies is slightly bigger with the cooperation than with out it; 28 vs. 26. This could be seen as an indication of how the network-centric access selection strategy is better able to exploit the network cooperation.

	Connected MNs	Norm. USI
Legacy - No Coop	67%	34
Network - No Coop	80%	60
Legacy - Coop	85%	70
Network - Coop	99%	98

Table 6.6: Connection and (norm.) USI statistics for connectivity sensitive users (200 mobile nodes) [PIII]

6.4 Operator satisfaction results

The main goal of the operator satisfaction calculations was to study what kinds of impacts the network cooperation would have when the used access selection algorithm is also able to exploit the cooperation. As according to the other experiment's results the network-centric access selection strategy was the best one to exploit the cooperation, it is used for the operator satisfaction measurements. The OSI calculations for both operators using the third P class weight set ($[-1, 0.25, 1, 1.4]$) and the network-centric strategy are presented in Table 6.7.

		USI	f_{RC}
Operator-1	No Coop	95.11	111.26
	Coop	492.00	137.86
Operator-2	No Coop	55.20	106.74
	Coop	474.26	132.92

Table 6.7: Average OSI calculations per user [PIV]

When we consider what is the impact of having the cooperation based on these operator satisfaction calculations, it can be said that the cooperation results in around 24% higher average f_{RC} per user for both operators. One important thing these results illustrate is that even if both operators have similar f_{RC} results, still operator-2 has clearly lower user satisfaction results than operator-1 especially without the cooperation. The difference is not so big when the cooperation is present. This is an example of how the cost revenue function and the overall USI are not directly comparable as explained in Section 4.2.

		USI	f_{RC}
Operator-1	No Coop	95.11	529.81
	Coop	492.00	656.46
Operator-2	No Coop	55.20	508.27
	Coop	474.26	632.94

Table 6.8: Average OSI calculations per user [PVI]

Table 6.8 shows another OSI calculations and compared to the first results presented in Table 6.7, there are huge differences between the absolute f_{RC} values. This is due to an error while calculating f_{RC} values, i.e., for some services, production costs were omitted and this resulted in higher f_{RC} values. In practice, this error is not visible in the USI calculations and since the conclusions and results presented in [PVI] [PVII] are not based on the absolute values this error has no impact.

6.5 Main findings

From a scalability point of view, the single operator simulation results presented in Section 6.1 clearly show that the two new access selection strategies outperform the legacy strategy as expected. The network-centric strategy does not fall so much behind the optimal solution and with a high network load the difference is only a few percent, which is a quite impressive performance considering that the LP is too time and resource consuming to be used in real-time and in real systems. The difference between the network-centric and the legacy strategies is fairly constant whereas the performance of the terminal-centric strategy approaches the network-centric one as the network load increases.

So clearly the terminal-centric strategy does not scale as well as the network-centric one; but what is the cost of this improved scalability? It is an increased amount of handovers and according to this metric, the performance of the terminal-centric strategy does not look so bad anymore. For the network-centric strategy, increased signalling and processing load to handle additional handovers could translate into an increased OPEX for an operator. The terminal-centric strategy yields only slightly more handovers compared to the legacy case, so by comparing the ratio of improved scalability and number of handovers, one could even argue that the the terminal-centric strategy is better than its competitors. However, as mentioned before, the access selection algorithm used for these experiments does not have a load balancing threshold implemented resulting in gratuitous handovers especially with lightly loaded networking conditions.

Additionally, the single operator simulation model does not include a handover model, which partially can skew the related results, because in the real world a handover always yields some computational cost. Therefore, the results should be seen more as indicative to show where we could potentially get with a powerful handover framework that would allow an operator to increase its effective network capacity with the appropriate handover decision algorithm resulting in additional revenues and lower relative OPEX due to a higher utilization rate. In addition, when the allowed offline time percentage for the active mobile nodes is required to be low, i.e., less than 5%, then the new access selection strategies provide substantial gains over the legacy one.

The simulation model for the multi-operator experiments includes many improvements compared its predecessor like an improved handover decision algorithm that is better suited for a multi-operator environment, a handover model and a load balancing threshold to tackle the "ping pong" effect. All these results support the earlier findings and now the network-centric strategy does not involve so many handovers compared to its rivals implying that the implemented load balancing threshold in the algorithm really works. The number of used intra-RAT handovers was approximately doubled compared to the legacy access selection strategy. However, the terminal-centric strategy now performed worse than in the single operator case in terms of a number of required inter-RAT handovers. And the network-centric strategy required almost twice as many inter-operator handovers than the terminal-centric case did. So considering both simulation setups, additional handovers are really required to optimize network utilization. Whether this is justifiable, it depends on how effectively handovers can be carried out including both horizontal and vertical ones.

The scalability results are consistent with the respective single operator results, i.e., two new access selection strategies do scale better when the network load increases. The disconnectivity results for 300 mobile nodes show a similar trend for the terminal-centric strategy that was already visible in the scalability results of the single operator case, i.e., in the beginning, the performance is close to the legacy case but once the load balancing and traffic off-loading really start to work, then its performance starts to approach the network-centric case. The network side decision making constraints are important for the exploitation of the network cooperation, which is also visible in the results, i.e., the difference between the network-centric strategy and its rivals is greater with the cooperation than without it. However, these results are only indicative of the potential technical benefits an enhanced handover strategy could produce. It should be also noted that the simulation results would be different if for instance different decision making constraint weights or different decision making constraints were used.

For the user satisfaction modelling and measuring, it is very important to make sure that a reasonable user happiness function is selected as showed clearly by the P class weight set comparison. The USI calculations show how large differences there can be between different weight sets. One important decision is whether disconnections ($P1$) should have a negative impact on the user happiness and if no negative impact is used, then the importance of the network cooperation is hardly visible in the USI calculations, i.e., only about 20% increase in the USI values (727 vs. 903). There can be situations and use cases where disconnections are not necessarily so drastic, e.g., the user is not a person, but sensor. Nevertheless when the user is human being, then it does matter whether (s)he is connected or not, which is also confirmed by the user survey presented in [PVII] [PVIII]. The user happiness function representing both connectivity and quality sensitivity results in the lowest and highest USI values and for this function, the presence of the network cooperation was most visible.

Non-linearity of the user happiness function (modeling the quality sensitivity) where the difference between weights $P2$ and $P3$ is bigger than the respective difference between weights $P4$ and $P3$ is also confirmed by the same user survey, i.e., the interviewed users valued connectivity and service availability in such way. However it is not clear how much this difference should be nor whether it is case sensitive. For instance, the expectations of a technically educated user could be different than a non-technically educated user, i.e., an engineer could either be more critical or more forgiving of disconnections than an average user that has no technical background.

The relationship between the network cooperation and the access selection strategies' ability to exploit the cooperation is shown in Table 6.9 where the changes of the normalized USI values are presented with the cooperation. The normalization is done so that the range of the normalized USI is $[-100, 100]$, where for instance the value 100 represents the maximum real USI value. The network-centric strategy has the biggest increments in its normalized USI values for both traffic profiles implying that the strategy gains more from the cooperation than its rivals. The terminal-centric and legacy cases are fairly close to each other, whereas the network-centric case is a clear winner, which was expected based on the technical single operator and multi-operator results and analysis.

	Connection	Connection& Quality
Legacy	+45	+53
Terminal	+48	+55
Network	+60	+65

Table 6.9: The cooperation and increased normalized USIs

The cooperation's existence increased more USI values than f_{RC} values, which was expected when considering the relation between these two indicators. In the OSI model, the cost revenue function (f_{RC}) represents the operator's revenue and naturally is something that operators try to maximize. The USI part then reacts on this, and in case of negative trend, the USI can have a negative impact on the cost revenue part in term of churn, as also indicated in [41]. However, it is unclear what is exactly the correlation between the USI and the cost revenue function. The USI is something that operators would not try to optimize as an independent parameter, thus it is part of their business model and optimization.

In addition, the cost revenue (f_{RC}) values are highly dependent on the used pricing and production cost models as the OSI results illustrate. Flat rate tariffs are especially challenging, since potentially additional traffic increases an operator's costs (assuming that each transferred data unit cost the same amount for the operator) but does not increase positive monetary flow. The cooperation increased the f_{RC} by about 20%, but there could be indirect impacts as well. For instance, a good reputation among customers and the resulted improved brand value could lead to a larger customer base and new business opportunities with other business players like advertisers.

Chapter 7

Concluding remarks

This dissertation proposes a new distributed access selection algorithm, which aims to make the access selection decision process more intelligent compared to the current practices by introducing a rich set of decision making information. This information captures both end-user and operator preferences. In order to adjust the importance of the end-user and operator preferences in the decision making, two new access selection strategies, the network-centric and terminal-centric, were developed. As their names imply, they weight decision making constraints differently, i.e., the network-centric weights the network constraints over the terminal ones. The proposed algorithm can exploit network cooperation when available. Network cooperation can be implemented and supported in different ways and the network composition [38] [16], which is one of the Ambient Networks project's main results, is one option to support a distributed access selection. This leads us to two questions: i) what kind of technical benefits the new distributed access selection algorithm provides and ii) what is the role of network cooperation in it?

In order to study these questions, the simulation model representing a distributed decision making in the access selection was developed (Chapter 5). This model was used for both single-operator and multi-operator scenarios and it includes a configurable network cooperation support option. The simulation experiments were carried out to measure how well different access selection methods perform. The simulation results and their analysis (Chapter 6) indicate that the network cooperation could be and in many cases is beneficial for both end users and operators due to increased coverage and improved service availability. However, how beneficial it is depends on the used access selection method as shown in the single operator (Section 6.1) and multi-operator (Section 6.2) results where the legacy strategy is not able to produce as much benefits as the new strategies when the network cooperation is taken into use. For instance, from an online time point of view, the legacy strategy results in about 10% worse connectivity results than the network-centric strategy. No matter whether the network cooperation is used or not, the new strategies (the network-centric and terminal-centric) result in clear technical benefits like increased connectivity time for end users and these are visible in the results.

In addition, the single operator measurements propose that the load balancing at the radio cell level could be improved if for instance the traffic type parameter is also taken into account in the access selection. Today, performance studies for the access and network selection algorithms and methods are typically based on realistic radio modelling where the access selection is based on quite heavily on the radio measurements and a tight network control [44] [76] [32] [82] [21] [26]. The proposed new algorithm and two access selection strategies do not follow these kinds of working assumptions, thus the starting point was to study a feasibility of new access selection methods. These methods use a rich set of information in the decision making including also end user's preferences. The decision making considers both terminal and network side constraints and differentiates them like for instance in [80].

It is more straightforward to measure the gained technical benefits than to try to assess how the satisfaction of different players is impacted. The simulation results show substantial gains in network utilization due to the ability to perform effective load balancing inside a radio access network and in between radio access networks. This can be also seen as an increased effective network capacity, which indicates a higher Return Of Investment (ROI) for the operators. The network-centric strategy was the clear winner, but its usage also has some drawbacks like an increased number of required handovers. However this strategy clearly shows potential behind the use of a richer set of information in the handover decision making and access selection.

When no additional cost for handovers is included, the network-centric strategy was quite close to the theoretical upper bound; the gap was only between 2% and 10%. On the other hand, the terminal-centric strategy was also able to produce substantial gains compared to the legacy case without requiring a significant amount of additional handovers. By adjusting decision making constraint weights, the algorithm behaviour can be changed and perhaps the best performance in practice could be found in between the terminal and network-centric strategies, since these two strategies represent more extreme cases of the full network or terminal control. Nevertheless, the results indicate the importance of having network cooperation when it is exploited properly. Also it is clear that the new access selection strategies provide technical benefits, which was one of the research questions.

Besides the technical evaluation of the new access selection algorithm and the access selection strategies, it is important to consider also non-technical aspects like the satisfaction of the related players like end users and operators. So how can the end-user and operator satisfaction related to the access selection be modelled and measured? There are many existing models representing user happiness and each model has its own metrics based on which the happiness is then measured. Most of these models use some form of Quality of Service (QoS) in measuring user happiness. This is also a valid starting point for user happiness modelling from the access selection point of view. In order to study these non-technical aspects, two new satisfaction models for end users and operators were developed.

For the satisfaction modelling and measuring, two new performance metrics were developed; the User Satisfaction Index (Chapter 3) and the Operator Satisfaction Indicator (Chapter 4). These metrics are used to study how the gained technical benefits translate into enhanced satisfaction of the end users and operators. It is important to differentiate between best-effort and real-time traffic models due to their different resource requirements compared to radio access networks. In a situation where a best-effort application session can provide an expected user experience it could be that a real-time application session may not be able to do so. This was one requirement for the design of the User Satisfaction Index (USI). The USI combines 3 different models (benefits, costs and success rate) used in [58], which naturally means some degree of simplification in the model. Nevertheless, the model is still accurate enough for the purposes it was designed for, i.e., to assess new access selection methods from the end-user happiness point of view.

In addition, the USI follows the findings from [91] resulting in non-linear user happiness function behaviour around the reference expectation level. It is difficult to estimate end users' expectations and this estimation could have a great impact on the satisfaction measurements. It is important that the used P class weights are selected properly since there are significant differences in the USI calculations depending on the used weights with the same measurement data. Considering the USI calculations from the technical measurements point of view, the calculations are compatible with the technical results. In other words, the user happiness model can successfully express the different user experiences. The network cooperation results in added user satisfaction according to the USI calculations.

The Operator Satisfaction Indicator (OSI) model is based on two factors; i) a happiness of all customers and ii) the relation between costs and revenue. The former is provided by the USI and for the latter, the cost revenue function is defined. These two factors represent different perspectives and do not directly relate. For instance, there can be big differences between the overall USIs, but the cost revenue function can produce almost identical values, or the other way around. The reason for this is that both the USI and OSI are designed to make measurements over a short time period like in simulation experiments.

Like the USI results, the OSI results also support the technical results and the calculations show that also operators receive more revenue because of increased connectivity, but an amount of extra revenue depends on the end-user pricing and other factors like market situation and cost structure. No matter whether the network cooperation is used or not, the new access selection strategies (the network and terminal-centric) result in clear added value for both end users and operators over the legacy one. This is visible in the technical measurements that show increased connectivity for end users, which translate into increased satisfaction index values for end users and this has a positive impact on the operator satisfaction indicator.

Network cooperation further amplifies these gained benefits when the access selection method can sufficiently exploit it and the legacy strategy benefits 33% less from the network cooperation than the network-centric one in terms of user satisfaction (USI). The lack of cooperation between operators is not always due to technical challenges, thus it could also be a result of the current practices in operator business or it could reflect the current market conditions such as a highly competitive market situation. For instance, because of tight vertical integration in the operators business landscape, it is just natural that national roaming is not typically supported.

Another important finding is that technical benefits and satisfaction benefits scale in a different way. For instance, if the difference between case-A and case-B is 30% better network utilization, this does not mean that the corresponding performance metrics are also increased by 30%. In other words, how the technical benefits translate into added satisfaction for the end users and operator is case sensitive and depends highly on the reference expectation level used and also the models used. Like the USI model, also the OSI model is able to express differences based on the technical results.

The results are encouraging and they support a current trend in the operating of mobile networks, i.e., when possible, users are moved to use cheaper accesses. This includes both off-loading the traffic from more expensive wide area coverage mobile networks to Wireless LAN (WLAN) or femto-cell hot spots and supporting a local breakout in the visited network. This trend is already visible in the current Mobile Network Operators (MNOs) business. For instance in Sweden two mobile operators have announced cooperation with local hot spot providers. Telia[1] and Tele2[2] are cooperating with The Cloud[3], which is Europe's largest independent Wireless Fidelity (WiFi) access provider. In the UK, O2[4] has announced a joint project with The Cloud to support off-loading.

Typically these types of off-loading scenarios require vertical handovers or a network re-attachment, which could result in a short connection break. The network side can help determining whether a mobile device is in a potential off-loading area. To develop a distributed access selection between a mobile device and network is not a trivial undertaking, at least not when considering a "true" generic mechanism that is not handset vendor or mobile operator specific. For such generic mechanism, standardization efforts in the relevant forums like 3rd Generation Partnership Project (3GPP) would be required.

With the Long-Term Evolution (LTE) standardization, an inter-working with non-3GPP networks is further extended introducing enhanced means to perform the access selection. The 3GPP specification [9] defines how trusted and non-trusted non-3GPP accesses could be supported by a 3GPP network including generic network discovery and selection rules that also apply in a multi-radio environment. So clearly this is a step in the right direction for exploiting different kinds of radio access technologies.

However, there are also other ways to gain progress in this area, i.e., an operator specific solution. This type of solution would be a proprietary solution working only within a specific operator's networks and its related business partners. In many market areas, like in UK or in north America, mobile operators sell their branded handsets that are tailored for them by the handset manufacturers. For instance, in north America there is Verizon [5] and correspondingly in the UK and Europe there are Vodafone[6] and T-mobile[7] to list a few. An example of such a proprietary solution to improve access selection by providing automatic usage of WiFi when possible, is the service AT&T™launched [99] for a new smart phone platform

---

[1] `http://www.telia.se/international`

[2] `http://www.tele2.com`

[3] `http://www.thecloud.net/About-us`

[4] `http://www.O2.com`

[5] `http://www.verizonwireless.com/b2c/index.html`

[6] `http://www.vodafone.co.uk`

[7] `http://www.t-mobile.com`

by Apple^{TM}. These types of services could act as pioneers by introducing new ways to manage and control radio access resources. This could encourage the deployment of such mechanisms as part of future architectural standardization of mobile networks. The standardization has an important role, since in a diverse market where multitude of non inter-operable proprietary solutions are deployed end users could be limited to using these advanced features only in certain networks. In such cases, the end users' freedom to use available accesses is limited.

There are also other reasons why the access selection practices should be further developed like a possible communication paradigm shift from the host based networking that we have today towards Information-Centric Networking (ICN) [37] [54] [36]. According to this paradigm, the focus is shifted from hosts to information implying that in such environments not only application session and traffic types are important, thus this differentiation should be more fine-grained.

In the Internet Engineering Task Force (IETF) [8] where Internet related standardization is mainly done, new ways of transporting information are actively studied in its the transport areas[9]. One very interesting activity in this area is the MultiPath Transmission Control Protocol (MPTCP) [10], which allows a session to be divided into multiple parallel flows similar to what is done by the Stream Control Transmission Protocol (SCTP) [69]. This together with multi-access capable mobile devices create yet new interesting possibilities for MNOs. So once the MPTCP or other similar transport protocols are widely deployed in the Internet, it is reasonable to expect that end users will be able to use them also with the mobile networks.

Traffic off-loading as discussed in this research could be "partial" in the sense that only some of the end users' transmission needs will be off-loaded to cheaper hot spot type of accesses like femto and WLAN hot spots. This partiality makes it possible that the main control (="control channel") that does not involve so much data transmission but which is crucial for an application session will be maintained in the wide range coverage accesses while a potential user is moving. In this way, even if a user suddenly moves out of the hot spot's coverage, the on-going application sessions could be better maintained and adapted, since the application sessions do not need to go through the session re-establishment phase, like bootstrapping a new session. These kinds of use scenarios require more

---

[8] `http://www.ietf.org`

[9] `http://datatracker.ietf.org/wg/`

[10] `http://datatracker.ietf.org/wg/mptcp/`

sophisticated ways of doing access selection where the indications both from the end user and from higher levels are taken into account, and the proposed access selection strategies are one step in that direction. One of the main challenges which still remains, is how to ensure that the increased complexity in the access selection does not interfere with usability while still providing enough flexibility.

References

[1] 3GPP TR 22.980 V9.0.0. 3GPP - Network composition feasibility study. 3GPP Technical Report, December 2009.

[2] 3GPP TR 36.913 V9.0.0. Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced) (Release 9). 3GPP Technical Report, December 2009.

[3] 3GPP TS 21.101 V5.14.0. Technical Specifications and Technical Reports for a UTRAN-based 3GPP system (Release 5). 3GPP Technical Standard, December 2009.

[4] 3GPP TS 21.101 V8.2.0. Technical Specifications and Technical Reports for a UTRAN-based 3GPP system (Release 8). 3GPP Technical Standard, December 2009.

[5] 3GPP TS 23.060 V10.2.0. General Packet Radio Service (GPRS); Service description; Stage 2 (Release 10). 3GPP Technical Standard, December 2010.

[6] 3GPP TS 23.234 V9.0.0. 3GPP - WLAN Interworking; System description (Release 9). 3GPP Technical Standard, December 2009.

[7] 3GPP TS 23.251 V10.0.0. Network Sharing; Architecture and Functional Description. 3GPP Technical Standard, December 2010.

[8] 3GPP TS 23.401 V10.1.0. General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access (Release 10). 3GPP Technical Standard, September 2010.

[9] 3GPP TS 23.402 V10.1.0. Architecture enhancements for non-3GPP accesses (Release 10). 3GPP Technical Standard, September 2010.

[10] 3GPP TS 24.302 V10.1.0. Access to the 3GPP Evolved Packet Core (EPC) via non-3GPP access networks; Stage 3 (Release 10). 3GPP Technical Standard, September 2010.

[11] 3GPP TS 25.855 V5.0.0. High Speed Downlink Packet Access; Overall UTRAN Description (Release 5). 3GPP Technical Standard, January 2001.

[12] 3GPP TS 43.318 V9.0.0. Generic Access Network (GAN); Stage 2 (Release 9). 3GPP Technical Standard, December 2009.

[13] 3GPP TS 44.318 V9.2.0. Generic Access Network (GAN); Mobile GAN interface layer 3 specification (Release 9). 3GPP Technical Standard, March 2010.

[14] 3GPP TS 50.059 V4.0.1. Enhanced Data rates for GSM Evolution (EDGE); Project scheduling and open issues for EDGE (Release 4). 3GPP Technical Standard, September 2001.

[15] N. Akhtar, J. Markendahl, and O. Queseth. Analysis of Signaling Load and Negotiation Complexity using Network Composition in Multi-Provider Business Environments. *Proc. of the 6th Conference on Telecommunication Techno-Economics, IEEE Computer Society*, pages 1–8, 2007.

[16] J. Andres-Colas and et al. Design of Composition Framework - FP6-CALL4-027662-AN P2/D3-G.1. Technical report, Ambient Networks project, November 2006.

[17] L. Badia, M. Lindstrom, J. Zander, and M. Zorzi. Demand and pricing effects on the radio resource allocation of multimedia communication systems. *Globecom 2003*, 2003.

[18] C. Bettstetter, H.-J. Vogel, and J. Eberspacher. GSM phase 2+ general packet radio service GPRS: Architecture, protocols, and air interface. *IEEE Communications Surveys & Tutorials, IEEE Computer Society*, 2:2–14, 1999.

[19] H. Bing, C. He, and L. Jiang. Performance analysis of vertical handover in a UMTS-WLAN integrated network. *Proceedings on 14th IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, Volume 1:187 – 191, 2003.

[20] R. Chakravorty, P. Vidales, K. Subramanian, I. Pratt, and J. Crowcroft. Performance issues with vertical handovers - experiences from GPRS cellular and WLAN hot-spots integration. *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications (PerCom)*, pages 155 – 164, 2004.

[21] A.P. da Silva, F. de S. Chave, V.A. de Sousa, R.A. de O. Neto, and F.R.P. Cavalcanti. Performance evaluation of access selection algorithms for VoIP on wireless multi-access networks. *International Telecommunications Symposium, IEEE Computer Society*, pages 663–667, 2006.

[22] R. Edell and P. Varaiya. Providing Internet Access: What we learn from the INDEX Trial. *IEEE Networks*, 1999.

[23] R J. Edell and et al. INDEX: A platform for determining how people value the quality of their Internet Access - project report #98-010P. Technical report, http://people.ischool.berkeley.edu/~hal/index-project/S98_010P.PDF, 1998. accessed March 2012.

[24] M. Eriksson, A. Furuskar, F. Kronestedt, C. Lindheimer, S. Mazur, J. Molno, and C. Tidestav. System overview and performance evaluation of GERAN-the GSM/EDGE Radio Access Network. *IEEE Wireless Communications and Networking Conference (WCNC), IEEE Computer Society*, 2:902–906, 1999.

[25] R. Ferrus, O. Sallent, and R. Agusti. Interworking in heterogeneous wireless networks: comprehensive framework and future trends. *IEEE Wireless Communications*, 2:22–31, 2010.

[26] X. Gelabert, J. Perez-Romero, O. Sallent, and R. Agusti. A Markovian Approach to Radio Access Technology Selection in Heterogeneous Multiaccess/Multiservice Wireless Networks. *IEEE Transactions on Mobile Computing, IEEE Computer Society*, 7:1257–1270, 2008.

[27] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas. LTE-advanced: next-generation wireless broadband technology [Invited Paper]. *Wireless Communications IEEE, IEEE Computer Society*, 17:10–22, 2009.

[28] J. Gozalvez and et al. User QoS-based Multi-Channel Assignment Schemes under Multimedia Traffic Conditions. *Proc. of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, 2007.

[29] GSM Association. Permanent Reference Document (PRD) AA.39:WLAN/GSM Roaming User Scenarios. GSM Association.

[30] GSM Association. Permanent Reference Document (PRD) SE.27:Services, Ease of Use, and Operator Considerations in Interworked WLAN-Cellular Systems. http://www.eduroam.hr/doc/se27.pdf.

[31] GSM Association. IR.34: Inter-Service Provider IP Backbone Guidelines. http://www.gsmworld.com/documents/ir3444.pdf, June 2008.

[32] A. Hasswa, N. Nasser, and H. Hassanein. Performance evaluation of a transport layer solution for seamless vertical mobility. *International Conference on Wireless Networks, Communications and Mobile Computing, IEEE Computer Society*, pages 576–581, 2005.

[33] D. Hollos, P. Pöyhönen, O. Strandberg, R. Aguero, K. Pentikousis, and 0. Blume. A study of handover strategies for mobile multiaccess Ambient Networks. *Proc. of the 16th IST Mobile and Wireless Communications Summit*, pages 1–5, 2007.

[34] IEEE Std 802.11. IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Standard, June 2007.

[35] IEEE Std 802.11. IEEE Standard for Local and Metropolitan Area Networks - Part 21: Media Independent Handover Services. IEEE Standard, January 2009.

[36] V. Jacobson and et al. Named Data Networking (NDN) Project - NDN-0001. Technical report, NDN project, October 2010. available online from http://www.named-data.net/, accessed March 2012.

[37] V. Jacobson and et al. Content-centric networking: Whitepaper Describing Future Assurable Global Networks. Technical report, Palo Alto Research Center, January 30, 2007. Presented in response to DARPA Assurable Global Networking RFI SN07-12.

[38] M. Johnsson, A. Schieder, R. Hancock, and et al. Final System Description - FP6-CALL4-027662-AN P2/D18-A.4. Technical report, Ambient Networks project, December 2007.

[39] I. Kadayif, M. Kandemir, N. Vijaykrishnan, and M. J. Irwin. An integer linear programming-based tool for wireless sensor networks. *J. Parallel Distrib. Comput.*, 65:247–260, 2005.

[40] M. Kassar, B. Kervella, and G. Pujolle. An overview of vertical handover decision strategies in heterogeneous wireless networks. *Computer Communications*, Volume 31, Issue 10:2607–2620, 2008.

[41] A. Lambrecht and B. Skiera. Paying too much and being happy about it: Existence, Causes, and consequences of tariff-choice biases. *Journal of Marketing Research*, XLIII, 2006.

[42] J. Lee and S. Bahk. Point of attachment selection in heterogeneous radio access technology environments. *The IEEE Symposium on Computers and Communications (ISCC)*, pages 873–878, 2010.

[43] C. Lindemann, M. Lohmann, and A. Thummler. A unified approach for improving QoS and provider revenue in 3G mobile networks. *ACM Journal on Special Topics in Mobile Networks and Applications (MONET)*, 2003.

[44] M. Liu, Z.-C. Li, X.-B. Guo, E. Dutkiewicz, and D.-K. Zhang. Performance Evaluation of Vertical Handoff Decision Algorithms in Heterogeneous Wireless Networks. *IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–5, 2006.

[45] Analysys Research Ltd. 3G Network Evolution from 2007 to 2012: HSPA+, LTE, WiMAX and femtocells. Technical report, Analysys Research Ltd, 2008.

[46] L. Ma, F. Yu, V. C. M. Leung, and T. Randhawa. A New Method to Support UMTS/WLAN Vertical Handover Using SCTP. *IEEE Wireless Communications*, pages 44–51, August 2004.

[47] J. Markendahl, H. Ensing, P. Karlsson, and M. Johnsson. Business opportunities and regulatory issues of Ambient Networking. *ITS Europe*, 2006.

[48] J. Markendahl and et al. Systems Evaluation Results - FP6-CALL4-027662-AN P2/D27-H.5. Technical report, Ambient Networks project, December 2007.

[49] J. Markendahl, P. Pöyhönen, and O. Strandberg. Performance Metrics for Analysis of Operator Benefits of Network Cooperation in Multi-Operator Business Scenarios. *The 19th Annual IEEE Int. Symposium*

on Personal, Indoor and Mobile Radio Communications (PIMRC'08), 2008.

[50] J. Markendahl, J. Werding, P. Pöyhönen, and O. Strandberg. Operator Cooperation as a Competitive Advantage for Provisioning of Low Cost High Capacity Mobile Broad band Services. *19th European Regional ITS Conference, Rome, Italy (ITS2008-Europe)*, 2008.

[51] D. Mulvey. HSPA. *Communications Engineer, IET JOURNALS, IEEE Computer Society*, 5:38–41, 2007.

[52] N. Niebert and et al. Ambient Networks Framework Architecture - IST-2002-507134-AN/WP1/D15. Technical report, Ambient Networks project, December 2005.

[53] Y. Nkansah-Gyekye and J. I. Agbinya. Vertical Handoff Decision Algorithms Using Fuzzy Logic. *First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06), IEEE Computer Society*, 2:309–313, 2006.

[54] B. Ohlman and et al. 4WARD Architecture and Design for the Future Internet - FP7-ICT-2007-1-216041-4WARD/D-6.1. Technical report, 4WARD project, 2008.

[55] O. Ormond, G.-M. Muntean, and J. Murphy. Network Selection Strategy in Heterogeneous Wireless Networks. *Proc. of the Information Technology and Telecommunications (ITT) Conference*, 2005.

[56] P. Faratin and D. Clark and P. Gilmore and S. Bauer and A. Berger and W. Lehr. Complexity of Internet Interconnections: Technology, Incentives and Implications for Policy. *Telecommunication Policy Research Conference (TPRC-07)*, 2007.

[57] K. Pentikousis, R. Aguero, J. Gebert, J. A. Galache, O. Blume, and P. Paakkonen. The Ambient Networks heterogeneous access selection architecture. *Proc. of the First Ambient Networks Workshop on Mobility, Multiaccess, and Network Management (M2NM)*, pages 49–54, 2007.

[58] O. Pohjola and K. Kilkki. Value-based methodology to analyze communication services. *The 5th Conference of Telecommunication, Media and Internet Techno-Economics (CTTE2006)*, 2006.

[59] P. Pöyhönen, D. Hollos, H. Tang, O. Blume, R. Aguero, and K. Pentikousis. Analysis of Load Dependency of Handover Strategies in Mobile Multiaccess Ambient Networks. *12th IEEE Symposium on Computers and Communications (ISCC)*, pages 15–20, 2007.

[60] P. Pöyhönen, J. Markendahl, and O. Strandberg. Business analysis of flexible roaming and operator cooperation in multi-provider markets - Impact on traffic load distribution and user experience. *18th European Regional ITS Conference*, 2007.

[61] P. Pöyhönen, J. Markendahl, and O. Strandberg. Impact of operator cooperation on traffic load distribution and user experience in Ambient Networks business scenarios. *The 6th Annual Global Mobility Roundtable 2007*, pages 56–64, 2007.

[62] P. Pöyhönen, J. Markendahl, and O. Strandberg. Analysis of benefits of operator cooperation using end-user and operator performance metrics. *17th Biennial Conference on "The Changing Structure of the Telecommunications Industry and the New Role of Regulation", Montreal, Canada*, 2008.

[63] P. Pöyhönen, J. Markendahl, and O. Strandberg. Analysis of Network Cooperation in Terms of Operator and User Satisfaction. *The 7th Conference of Telecommunication, Media and Internet Techno-Economics (CTTE2008)*, 2008.

[64] P. Pöyhönen, J. Markendahl, and O. Strandberg. Analysis of User Experience of Access Selection in Multi-Operator Environments. *Proc. of the third International Conference on Systems and Networks Communications (ICSNC), IEEE Computer Society*, pages 27–32, 2008.

[65] P. Pöyhönen, J. Markendahl, O. Strandberg, J. Tuononen, and M. Johnsson. Analysis of Enhanced Access Selection Methods and End-User Perception in Multi-operator Environment. *International Journal On Advances in Intelligent Systems*, 2:107–125, 2009.

[66] P. Pöyhönen, O. Strandberg, J. Markendahl, and J. Laganier. Business Implications of Composition Framework in Ambient Networks. *Helsinki Mobility Roundtable, SPROUTS: Working Papers on Information Systems, ISSN 1535-6078*, 2006.

[67] P. Pöyhönen, J. Tuononen, H. Tang, and O. Strandberg. Study of Handover Strategies for Multi-Service and Multi-Operator Ambient Networks. *The Second International Conference on Communications*

and Networking in China (CHINACOM'07), IEEE Computer Society, pages 755–762, 2007.

[68] M. Prytz, J. Gebert, and et al. Multi-Access and ARI, Design and Initial Specification - FP6-CALL4-027662-AN P2/D02-C.1. Technical report, Ambient Networks project, September 2006.

[69] R. Stewart and et al. Stream Control Transmission Protocol. IETF Request for Comments: 4960, September 2007.

[70] J. Sachs. A generic link layer for future generation wireless networking. *Proc. of the IEEE International Conference on Communications (ICC)*, 2:834–838, 2003.

[71] Apostolis K. Salkintzis. Interworking Techniques and Architectures for WLAN/3G Integration Toward 4G Mobile Data Networks. *IEEE Wireless Communications*, pages 50–61, 2006.

[72] E. Sauerwein, F. Bailom, K. Matzler, and H. Hinterhuber. The Kano Model: How To Delight Your Customers. *Preprints Volume I of the IX. International Working Seminar on Production Economics*, pages 313–327, 1996.

[73] M. Sawahashi, Y. Kishiyama, H. Taoka, M. Tanno, and T. Nakamura. Broadband radio access: LTE and LTE-advanced. *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), IEEE Computer Society*, pages 224–227, 2009.

[74] A. Schieder, J. Zander, and R. Hancock. *Ambient Networks: Cooperative Mobile Networking for the Wireless World.* Wiley, ISBN: 978-0-470-51092-6, 2007.

[75] W. Shen and Q.-A. Zeng. A Novel Decision Strategy of Vertical Handoff in Overlay Wireless Networks. *Fifth IEEE International Symposium on Network Computing and Applications (NCA 2006)*, pages 227–230, 2006.

[76] T. Shu, M. Liu, and Z. Li. A performance evaluation model for RSS-based vertical handoff algorithms. *IEEE Symposium on Computers and Communications (ISCC)*, pages 271–276, 2009.

[77] Q. Y. Song and A. Jamalipour. A Network Selection Mechanism for Next Generation Networks. *IEEE International Conference on Communications (ICC), IEEE Computer Society*, 2:1418–1422, 2005.

[78] Q. Y. Song and A. Jamalipour. Vertical Handover Decision in an Enhanced Media Independent Handover Framework. *Proc. of the IEEE Wireless Communications and Networking Conference (WCNC), IEEE Computer Society*, pages 2693–2698, 2008.

[79] E. Stevens-Navarro, U. Pineda-Rico, and J. Acosta-ElÃas. Vertical Handover in beyond Third Generation (B3G) Wireless Networks. *International Journal of Future Generation Communication and Networking*, Volume 1, No 1:51–58, 2008.

[80] E. Stevens-Navarro and V.W.S Yuxia Lin Wong. An MDP-Based Vertical Handoff Decision Algorithm for Heterogeneous Wireless Networks. *IEEE Transactions on Vehicular Technology (TVT), IEEE Computer Society*, 57:1243–1254, 2008.

[81] O. Strandberg, J. Markendahl, and P. Pöyhönen. Connectivity And Access Provisioning With Ambient Networks. *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), IEEE Computer Society*, pages 1–5, 2007.

[82] M.Z.A. Syuhada, I. Mahamod, and W.A.W.N.S. Firuz. Performance Evaluation of Vertical Handoff in Fourth Generation (4G) Networks Model. *6th National Conference on Telecommunication Technologies and 2nd Malaysia Conference on Photonics (NCTT-MCP), IEEE Computer Society*, pages 392–398, 2008.

[83] H. Tang, J. Eisl, and et al. Mobility Support: Design and Specification - FP6-CALL4-027662-AN P2/D9-B.1. Technical report, Ambient Networks project, December 2006.

[84] H. Tang, J. Gebert, and et al. Multi-Access System Design and Specification - FP6-CALL4-027662-AN P2/D15-C.2. Technical report, Ambient Networks project, December 2007.

[85] The Advanced Resource Management Solutions for Future All IP Heterogeneous Mobile Radio Environments (AROMA) project website. http://www.aroma-ist.upc.edu/. accessed March 2012.

[86] The Ambient Networks (AN) project in Wikipedia. http://en.wikipedia.org/wiki/Ambient_network. accessed March 2012.

[87] The Ambient Networks (AN) project website. http://www.ambient-networks.org.

[88] The Evolutionary Strategies for Radio Resource Management in Cellular Heterogeneous Networks (EVEREST) project website. http://www.everest-ist.upc.es/. accessed March 2012.

[89] The IEEE 802.21 Working Group website - Media Independent Handover Services. http:/www.ieee802.org/21/. accessed March 2012.

[90] Nishith D. Tripathi, Jeffrey H. Reed, and Hugh F. Vanl. Handoff in Cellular Systems. *IEEE Personal Communications*, 1998.

[91] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.

[92] D. Walder. Kano's model for understanding customer-defined quality. *Center For Quality of Management Journal*, 39:65–69, 1993.

[93] W. Wu, N. Banerjee, K. Basu, and S. K. Das. SIP Based Vertical Handoff between WWANs and WLANs. *IEEE Wireless Communications*, pages 66–72, June 2005.

[94] B. Xing and N. Venkatasubramanian. Multi-constraint dynamic access selection in always best connected networks. *The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, IEEE Computer Society*, pages 56–64, 2005.

[95] R. Yallapragada, V. Kripalani, and A. Kripalani. EDGE: a technology assessment. *IEEE International Conference on Personal Wireless Communications, IEEE Computer Society*, pages 35–40, 2002.

[96] W. Zhang, J. Jaehnert, and K. Dolzer. Design and evaluation of a handover decision strategy for 4th generation mobile networks. *The 57th IEEE Semiannual Vehicular Technology Conference (VTC 2003-Spring)*, 3:1969–1973, 2003.

[97] F. Zhu and J. McNair. Optimizations for vertical handoff decision algorithms. *The IEEE Wireless Communications and Networking Conference (WCNC)*, 2:867–872, 2004.

[98] F. Zhu and J. McNair. Multiservice vertical handoff decision algorithms. *EURASIP Journal on Wireless Communications and Networking*, Volume 2006, Issue 2:52–52, 2006.

[99] C. Ziegler. AT&T's free hotspot access finally useful with auto-connection in iPhone OS 3.0.

http://www.engadget.com/2009/06/17/atandts-free-hotspot-access-finally-useful-with-auto-connection-i/, June 2007. accessed March 2012.