Optimization of Vertical Handover Decision Processes for Fourth Generation Heterogeneous Wireless Networks

A thesis submitted for the degree of
Doctor of Philosophy

by

Xiaohuan Yan

Department of Electrical and Computer Systems Engineering
Monash University
Australia

September 2010

Abstract

This thesis presents a vertical handover decision (VHD) scheme for optimizing the efficiency of vertical handover processes in the Fourth Generation (4G) heterogeneous wireless networks. The scheme consists of three closely integrated modules: Handover necessity estimation, handover target selection, and handover triggering condition estimation. Handover necessity estimation module determines whether a handover is necessary to an available network. Handover target selection module chooses the "best" network among the available candidates based on a set of criteria. Finally, handover triggering condition estimation module determines the right moment to initiate a handover out of the currently connected network.

4G wireless networks are expected to support mechanisms for tight integration and cooperation of divergent access network technologies. In such networks of heterogeneous nature, roaming users will experience frequent handovers across network boundaries. Thus, to ensure seamless roaming and efficient resource usage over dissimilar networks, intelligent VHD algorithms need to be used extensively. The research project presented in this thesis report focuses on this problem and provides an optimized VHD scheme, which minimizes the handover failures, unnecessary handovers and connection breakdowns whilst maintaining users' satisfaction at high levels. In addition, the scheme also provides mechanisms for mobile applications to control the tradeoff between the usage of the preferred access network and number of handovers or connection breakdowns.

Simulation based performance evaluations demonstrate that the scheme reduces the number of handover failures, unnecessary handovers and connection breakdowns by up to 80%, 70% and 70%, respectively. They also show an increase of up to 50% in the satisfaction level of users.

Publications Related to the Study Presented in This Thesis

Journal Articles

X. Yan, Y. A. Şekercioğlu, and S. Narayanan, "A survey of vertical handover decision algorithms in fourth generation heterogeneous wireless networks," *Computer Networks*, 54(11):1848-1863, August 2010.

X. Yan, N. Mani, and Y. A. Şekercioğlu, "A traveling distance prediction based method to minimize unnecessary handovers from cellular networks to WLANs," *IEEE Communications Letters*, 12(1):14-16, January 2008.

X. Yan, Y. A. Şekercioğlu, and S. Narayanan, "A probability based handover triggering condition estimation method for WLAN usage optimization," *Wireless Communications and Mobile Computing* (submitted).

Conference Papers

X. Yan, Y. A. Şekercioğlu, and N. Mani, "A method for minimizing unnecessary handovers in heterogeneous wireless networks," in *Proceedings of the 2008 International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'08),*, pages 1-5, Newport Beach, CA, USA, June 2008.

Declaration

I declare that, to the best of my knowledge, the research described herein is original except where the work of others is indicated and acknowledged, and that the thesis has not, in whole or in part, been submitted for any other degree at this or any other university.

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Xiaohuan Yan
Melbourne
June 2010

Acknowledgments

I owe my deepest gratitude to my supervisors, Dr. Ahmet Şekercioğlu, Professor Sathya Narayanan, and Professor Jean Armstrong. This work would not have been possible unless the support and guidance of them. Their enthusiasm and knowledge in the area made for an interesting and challenging project. A special thanks goes to Dr. Ahmet Şekercioğlu for his insightful advices on the project, and continued patience during the write up stage which greatly improved my technical writing skills.

I also would like to thank the staff of ECSE for their support during my studies at Monash University. My sincere appreciation goes to the office and administration staff for ensuring the countless forms and paperwork were correctly completed and filed.

I am deeply indebted to my friends, in particular Aisha Qi, Ke Liu, Xiaofang Zha, Himal Suraweera, Alan Guo, Joseph Violi, Martin Firus, Liang Du, Sarangi Dissanayake, Dinesh Segaran, Xia Li, Ke Ma, Yang Yang, Vajira Amaratunga, Sophia Kaplantzis, Han Liu, Xiwu Liu and Guanlin Zhang, for their help with my research, and more for their great friendship and encouragement during my studies.

Lastly and most importantly, I wish to thank my parents for their continued love and support. To my parents, I owe everything I have achieved in my life.

Contents

List of Figures

List of Tables

List of Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
AIPN	All-IP Network
AMPS	Advanced Mobile Phone System
ANN	Artificial Neural Networks
AP	Access Point
ASST	Application Signal Strength Threshold
BRAIN	Broadband Radio Access for IP-Based Networks
BS	Base Station
CAC	Call Admission Control
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CFC	Cost Factor Calculation
DECT	Digital Enhanced Cordless Telecommunications
EC	European Commission
EDGE	Enhanced Data rates for GSM Evolution
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HNE	Handover Necessity Estimation
HTCE	Handover Triggering Condition Estimation
HTS	Handover Target Selection
iDEN	Integrated Digital Enhanced Network
IETF	Internet Engineering Task Force
IMT-2000	International Mobile Telecommunications-2000
IP	Internet Protocol
IST	Information Society Technologies
ITU	International Telecommunication Union
LTE	Long Term Evolution
MAC	Medium Access Control

MICS	Media Independent Command Service
MIES	Media Independent Event Service
MIH	Media Independent Handover
MIHF	Media Independent Handover Function
MIIS	Media Independent Information Service
MIMO	Multiple-Input Multiple-Output
MIRAI	Multimedia Integrated Network by Radio Access Innovation
MT	Mobile Terminal
OFDM	Frequency-Division Multiplexing
PDC	Personal Digital Cellular
PDF	Probability Density Function
PEV	Performance Evaluation Values
PoA	Point of Attachment
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RSS	Received Signal Strength
RTMI	Radio Telefono Mobile Integrato
SAP	Service Access Point
SINR	Signal to Interference and Noise Ratio
TACS	Total Access Communications System
TDMA	Time Division Multiple Access
UMTS	Universal Mobile Telecommunications System
VHD	Vertical Handover Decision
VHDS	Vertical Handover Decision System
WCDMA	Wideband Code Division Multiple Access
WD	Weights Distribution
WDP	Wrong Decision Probability
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WWAN	Wireless Wide Area Network

List of Symbols

β	Path loss coefficient
Γ	dB gap between the uncoded Quadrature Amplitude Modulation and channel capacity
γ_{AP}	SINR at the MT when associated with WLAN
γ_{BS}	SINR at the MT when associated with WCDMA
θ_i	Entry angle of the WLAN cell coverage
θ_o	Exit angle of the WLAN cell coverage
ρ	Traffic load of the network
σ	Standard deviation of the shadow fading before averaging RSS
σ_a	Standard deviation of the shadow fading after averaging RSS
τ	WLAN to 3G handover delay
ASST	Application signal strength threshold
B	A point at which the MT reaches the boundary of the WLAN
B_i	Available bandwidth of network i
b_i, b_j	Available bandwidth of network i and j
$C(\cdot)$	Cost function
C_i	Cost factor
C_s^n	Per-service cost of network n
d	Side length of the WLAN cell
d_{ref}	Distance between the AP and a reference point
d_R	Half length of the trajectory segment inside the outer circle of the WLAN cell
d_r	Half length of the trajectory segment inside the inner circle of the WLAN cell
$E_{s,j}^n$	Network elimination factor of parameter j for service s provided by network n
$\text{EL}[k]$	Lifetime metric at time instant k
e_r	Averaged estimation error ratio
$F(\cdot)$	Cumulative distribution function
$f(\cdot)$	Probability density function
i_H	Factor of high importance level
i_L	Factor of low importance level
i_M	Factor of medium importance level

L_{BA}	Shortest distance between the point at which handover is initiated and WLAN boundary
l_{OP_i}	Distance between the AP and the entry point of the WLAN cell coverage
l_{OS}	Distance between the AP and where the MT takes an RSS sample
l_{OS_1}	Distance between AP location O and sampling point S_1
l_{OS_2}	Distance between AP location O and sampling point S_2
l_P	Distance between the AP and point P
l_{P_iM}	Distance between entry point P_i and middle point M
l_{S_1M}	Distance between sampling point S_1 and middle point M
$l_{S_1S_2}$	Distance between sampling points S_1 and S_2
l_{S_2M}	Distance between sampling point S_2 and middle point M
M	Middle point of the traveling trajectory
M_i	Monetary cost per minute of network i
N	Number of RSS samples
n_H	Number of high importance levels designated by the user
n_L	Number of low importance levels designated by the user
n_M	Number of medium importance levels designated by the user
n_N	Number of none importance levels designated by the user
O	AP location
P_{Tx}	Transmit power of the WLAN AP
PL_{ref}	Path loss at the reference point
P_b	Connection breakdown probability
P_f	Probability of a handover failure
P_i	Entry point of the WLAN cell coverage
P_o	Exit point of the WLAN cell coverage
P_r	Unnecessary handover probability
P_u	Probability of an unnecessary handover
p_w	Battery power level of the MT
$Q^n_{s,j}$	Normalized QoS provided by network n for parameter j and service s
R	Radius of the WLAN cell
$\overline{\text{RSS}}[k]$	Average of RSS at time instant k
RSS_{min}	Minimum level of the RSS required for the MT to communicate with an AP
RSS_B	RSS threshold for triggering
RSS_P	RSS at point P
RSS_{P_i}	RSS at the entry point P_i
RSS_S	RSS at sampling point S
RSS_{S_1}	RSS at sampling point S_1
RSS_{S_2}	RSS at sampling point S_2
r	Radius of the inner circle of the WLAN cell
$S[k]$	RSS change rate
S_{dth}	Dynamic RSS threshold
S_i	Security level of network i

S_{int}	Sampling interval
T_{WLAN}	Time threshold for handover from the cellular network to the WLAN
$T1$	Time threshold for minimizing handover failures from WLANs to cellular networks
$T2$	Time threshold for minimizing unnecessary handovers from WLANs to cellular networks
t_B	Time at boundary point B
t_{P_i}	Time at which the MT enters the WLAN cell coverage
t_r	Estimated remaining traveling time inside the WLAN
t_S	Time at which the RSS sample is taken
t_{S_1}	Time at sampling points S_1
t_{S_2}	Time at sampling points S_2
t_{WLAN}	Estimated traveling time inside the WLAN
v	Speed of the MT
W_{av}	Moving average window size
$W^n_{s,j}$	Weight of the QoS parameter provided by network n for parameter j and service s
w_B	Weight factor of available bandwidth
w_{i_H}	Weight factors of high importance level
w_{i_L}	Weight factors of low importance level
w_{i_M}	Weight factors of medium importance level
w_{i_N}	Weight factors of none importance level
w_M	Weight factor of monetary cost
W_n	Weight of the sample taken at the end of the $(k-n)^{\text{th}}$ interval
w_P	Weight factor of power consumption
w_S	Weight factor of security
X_σ	A Gaussian distributed random variable with a mean of zero and a standard deviation σ

Chapter 1

Introduction

Growing consumer demand for access to communication services anywhere and anytime is driving an accelerated technological development towards the integration of various wireless access technologies. Such integration combines islands of access networks into a seamless system, referred to as Fourth Generation (4G) wireless systems [AXM04, CYA10]. 4G wireless systems will provide significantly higher data rates, offer a variety of services and applications previously not possible due to bandwidth limitations, and allow global roaming among a diverse range of mobile access networks [DSVK07, NVAGD07, GB06, ZK03, HY03, VJ01].

In a typical 4G networking scenario, handsets or Mobile Terminals (MTs) with multiple interfaces will be able to choose the most appropriate access link among the available alternatives. These access links include IEEE 802.11 Wireless Local Area Network (WLAN) access [DY05], IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMAX) [WGS+08], satellite systems [Bea84] and Bluetooth [SGR02], in addition to the traditional cellular telephony networks. For a satisfactory user experience, MTs must be able to seamlessly transfer to the "best" access link among all available candidates with no perceivable interruption to an ongoing voice or video conversation. Such ability to handover between heterogeneous networks is referred to as vertical handovers [MK00]. As an important step towards achieving this objective, the emerging IEEE 802.21 standard creates a framework to support protocols for enabling seamless vertical handovers [TOF+09]. IEEE 802.21 provides only the overall framework, leaving the implementation of the actual algorithms to the engineers designing the system. Therefore, it is essential to develop efficient vertical handover decision (VHD) algorithms to ensure the success of this new framework.

The primary focus of this thesis is to develop a VHD scheme to optimize the efficiency of vertical handovers in heterogeneous 4G wireless networks.This chapter begins with Section 1.1 introducing the evolution of wireless communications and features of 4G wireless systems, followed by Section 1.2 which provides a brief overview of IEEE 802.21 Media Independent Handover (MIH) and explains why VHD algorithms are essential components of MIH. Then, Section 1.3 identifies open issues in existing VHD algorithms, and enumerates objectives and contributions of the research presented in this thesis. Section 1.4 provides an outline of the thesis structure.

1.1 Towards 'Always On, Always Best Connected' Communications

This section describes the development and features of 4G wireless systems. The evolution of wireless communications, and the definition of 4G and its features are highlighted. The key challenges in achieving 4G are also presented.

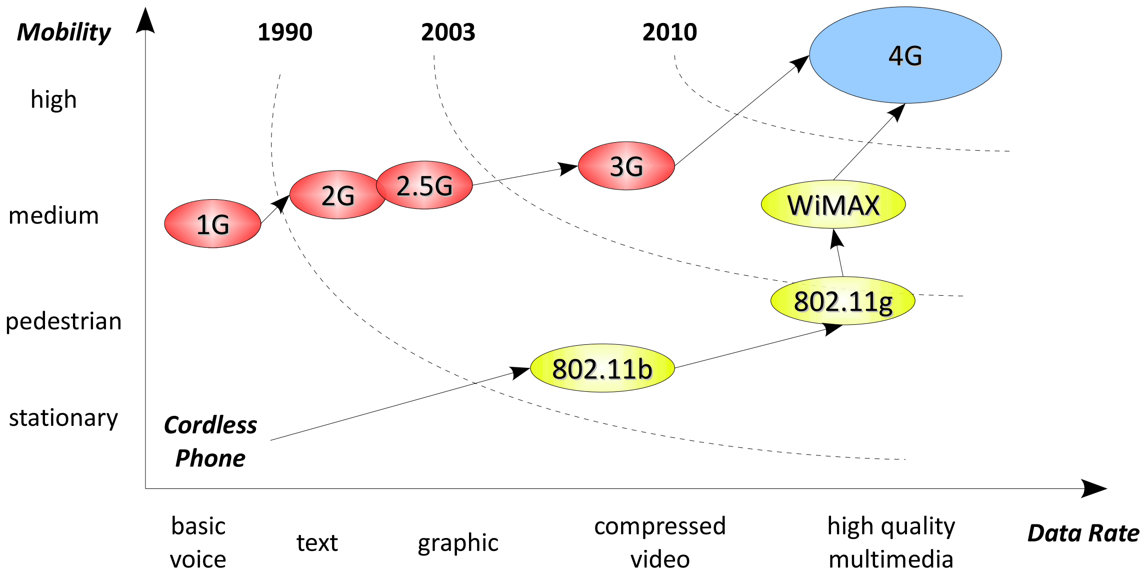

Figure 1.1: Evolution of wireless communications. From 1G to 4G, the data rate and user mobility have both increased.

1.1.1 Evolution of Wireless Communications

Today, communication technologies have become an integral part of people's daily life and the wireless communication market has grown rapidly. Driven by the increasing demands of the market, wireless communication technologies have evolved from the first to third generation and are moving towards 4G, as illustrated in Figure 1.1. Table 1.1 summarizes the development of wireless communications up to the Third Generation (3G), offering the properties of each generation by comparing the driving technology, representative standard, radio frequency, bandwidth, multi-address technique, core networks and service type.

Wireless communications began in the early 1980s, when the First Generation (1G) wireless telecommunication technology provided voice transmissions using frequencies around 900 MHz and analogue modulation. Different 1G standards were used in various countries, such as Advanced Mobile Phone System (AMPS) [You79] which was used in the United States, Total Access Communications System (TACS) [Tac02] in the United Kingdom, C-450 [Gol05] in Germany, Portugal

Generation / Property	1G	2G	2.5G	3G
Driving technology	Analogue signal processing	Digital signal processing	Packet switching	Intelligent signal processing
Representative standard	AMPS, TACS	GSM, I-Mode	GPRS, TDMA, HSCSD, EDGE	IMT-2000 (UMTS, WCDMA, CDMA2000)
Radio frequency (HZ)	400M – 800M	800M – 900M, 1800M – 1900M	800M – 900M, 1800M – 1900M	2G
Bandwidth (bps)	2.4K – 30K	9.6K – 14.4K	171K – 384K	2M – 5M
Multi-address technique	FDMA	TDMA, CDMA	TDMA, CDMA	CDMA
Core network	Telecom networks	Telecom networks	Telecom networks	Telecom networks, some IP networks
Service type	Voice	Voice, short message service	Data service	Voice, data, some multimedia

Table 1.1: Properties of different generations of wireless/mobile communications technologies (properties of 4G networks are separately provided in Section 1.1.2).

and South Africa, and finally Radiocom 2000 (RC2000) [Gol05] which was used in France.

The Second Generation (2G) of wireless networks that appeared in 1991 were based on low-band digital data signaling. Depending on the type of multiplexing involved, 2G technologies were categorized into Time Division Multiple Access (TDMA) based and Code Division Multiple Access (CDMA) based standards [DB98]. The main 2G standards were: TDMA based Global System for Mobile Communications (GSM) which was used worldwide [Ada01], integrated Digital Enhanced Network (iDEN) in the United States and Canada [Ada01], Interim Standard 136 (IS-136) in North and South America [Ada01], Personal Digital Cellular (PDC) in Japan [Ada01], and CDMA based IS-95 which was used in the Americas and parts of Asia [Ada01].

After 2G and before the 3G, a stepping-stone technology called Second and a Half Generation (2.5G) was used to describe 2G systems that had implemented a packet switched domain in addition to the circuit switched domain. 2.5G provided some of the benefits of 3G (since it was packet-switched) and could use some of the existing 2G infrastructure in GSM and CDMA networks [Mis04]. Its standards included General Packet Radio Service (GPRS) for GSM which was used worldwide [Mis04], and Enhanced Data rates for GSM Evolution (EDGE) which was used worldwide except for Japan and South Korea (EDGE was also considered as a 2.75G standard) [Mis04].

Since 2003, mobile telecommunications networks have been upgraded to use 3G technologies, also known as International Mobile Telecommunications-2000 (IMT-2000), defined by the International Telecommunication Union (ITU) [Cal97]. Compared to previous generations, 3G enables network operators to offer users greater bandwidth, security and reliability, and a wider range of more advanced services. 3G standards include CDMA2000 [KLR02] and Universal Mobile Telecommunications System (UMTS) [Ric00] which are used worldwide, and Digital Enhanced Cordless Telecommunications (DECT) in Europe and the United States [SBB+06].

The main disadvantage of current 3G networks is the high cost to both network operators and end users. For example, Base Stations (BSs) and cellular infrastructure need to be upgraded, different handsets are required, and very high spectrum-license costs need to be spent [Gar02].

Future wireless networks are expected to provide users with convenient global information access capabilities and personalized multimedia wireless communication services [LCCS05]. Growing interest in 4G networks is leading to a convergence of various wireless network technologies. Recently ratified IEEE 802.21 MIH standard [TOF+09] aims to support seamless roaming among a variety of wireless access network technologies, including GSM, UMTS, WiMAX, Bluetooth and WLAN, through different handover mechanisms.

Some of the world's leading carriers have already started working towards 4G. In January 2009, Clearwire and Intel collaborated to produce the world's first 4G network, called "Clear", in Portland, Oregon, U.S.A. [Lyl09]. WiMAX was adopted to allow consumers and businesses to "wireless connection anywhere in Portland at true broadband speeds" [Lyl09]. In addition, major mobile carriers in the United States such as AT&T and Verizon Wireless [Wil09], and several worldwide carriers are planning to convert their networks to 4G using another standard, 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE), a successor of UMTS [DPSB08]. While it is thought by some that LTE will be the standard adopted by 80 percent of the carriers in the world, the deployment of LTE will not be fully utilized until 2012 [Hig08].

1.1.2 4G Wireless Networks

Service providers, researchers and engineers have different views of 4G. These include the following standpoints:

- DoCoMo introduced the concept of MAGIC for the vision of 4G [Mur99]: Mobile multimedia; Anytime, anywhere, anyone; Global roaming support; Integrated wireless solution; and Customized personal service. MAGIC is particularly for public systems and it treats 4G as an extension of 3G cellular service.
- European Commission (EC) presented a perspective of 4G focusing on the seamless service across a multitude of wireless networks, and the optimum delivery via the available network. Further discussion provided continuous promotion around 4G concepts [Wil99, Huo99], e.g. private systems

and ad-hoc networks, optimal resource utilization, multiple radio interfaces, WLAN use, standards for interoperability, etc.

- A broader perspective of 4G was proposed in [Per00], according to which 4G will encompass all systems from public to private, operator-driven to ad hoc, broadband to personal and 2G to 3G. The focus of this paper was only on personalized services.

- Technical perspectives of 4G were presented in [SSH01]. A 4G feature framework was proposed based on the key concept of integration. Targets in the framework included users, terminals, networks and applications.

Despite the fact that no standards body has explicitly defined or agreed upon exactly what 4G will be, the key features of 4G remain common to researchers and are summarized as follows [HY03]:

1. **4G will provide high data rate at low transmission cost.**

   Along with telecommunications services, 4G systems will also provide data and multimedia services. To support multimedia services, it is necessary to have high data rate and reliable systems. The data rate of 4G is expected to be 10 times higher than 3G, with about 20 Mbps bandwidth, and peak bit rate up to 100 Mbps for high mobility and 1 Gbps in hot-spots [MOYU05]. Furthermore, it is necessary to keep the service cost-effective. A low per-bit transmission cost will be maintained in 4G.

2. **4G will be an all-Internet-Protocol based network providing anytime and anywhere communications.**

   Existing wireless systems can be classified into two types: Internet Protocol (IP) based and non-IP based. Many non-IP based systems are optimized for voice delivery (e.g. GSM, CDMA2000 and UMTS). By contrast, IP based systems are usually optimized for data services (e.g. 802.11 WLAN). 4G networks will integrate these two systems and have a structure based on all-IP. Using this system, IP packets will be able to traverse distinctive access networks connected to an IP based backbone network without any protocol conversion [KJC+03, CLB07]. Hence all-IP based heterogeneous networks will enable users to use any system anytime and anywhere. Users carrying an integrated terminal will be able to use a wide range of applications provided by multiple wireless networks.

3. **4G will always be connected to the best network.**

   4G systems will be based on a heterogeneous infrastructure comprising different wireless access systems. These systems will complement each other for different service requirements and radio environments. 4G mobile users will benefit from seamless mobility and ubiquitous access to the most efficient combination of available access systems in an “always best connected” mode [GAM05].

4. **4G will provide personalization.**

   When 4G services are launched, users will be expected from widely different locations, occupations and economic classes. In order to cater to the demands of these diverse users, service providers will need to design personalized and customized services. To achieve this, [FFF+06] defined 4G technology from the user's perspective. A user-centric methodology that considers the user as the “cornerstone” of the design was adopted.

1.1.3 4G Challenges and Development

Numerous problems need to be tackled to achieve 4G and a number of researchers have been working on these problems. In this section, a brief discussion on 4G challenges and development is given. The challenges are categorized into three groups: network systems, mobile terminals and services.

Network Systems

In 4G, various wireless network systems will coexist and interwork. Challenges reside in such network systems include the design of integrated network infrastructure, support for seamless terminal mobility and provision of Quality of Service (QoS). They are discussed below.

1. **Integrated Network Infrastructure:** In 4G, more advances in the network infrastructure are needed to provide seamless integration of heterogeneous wireless systems. As mentioned in Section 1.1.2, 4G will be based on an all-IP network infrastructure.

   Several solutions have been proposed to integrate heterogeneous wireless systems based on an IP network. Examples of these solutions are the All-IP Network (AIPN) proposed by 3GPP [PD02], Broadband Radio Access for IP-Based Networks (BRAIN) proposed by Information Society Technologies (IST) [BRA01], and Multimedia Integrated Network by Radio Access Innovation (MIRAI) proposed in Japan [WMH02]. In the following paragraphs the main characteristics of these architectures are briefly introduced.

   AIPN in the 3GPP standard is based on the GPRS protocol that was developed to provide packet services to the GSM [PD02]. The 3GPP access network is interfaced to the core network by a serving GPRS gateway node. In

2005, 3GPP published AIPN Release 7, which was the foundation of higher level protocols such as LTE. In 2008, Release 8 specification was finalized. The standard has been complete enough that hardware designers have been designing chipsets, test equipment and base stations for some time [Eri09].

The BRAIN project is an Information Society Technologies (IST) program [BRA01]. The BRAIN network architecture consists of a BRAIN Access Network (BAN), BRAIN Mobility Gateways (BMGs), a BRAIN Access Router (BAR), and an IP-based core network. The network components were imported from standard Internet Engineering Task Force (IETF) protocols to facilitate network evolution and flexibility. The access network is based on IP, and the access router interfaces the mobile node and access network. The gateway is placed between the access network and the core network.

MIRAI is a Japanese national project under the e-Japan plan for seamless integration of heterogeneous wireless systems [WMH02]. MIRAI architecture is composed of four major building blocks: a mobile host, Radio Access Networks (RANs), a Common Core Network (CCN), and an external IP network. CCN contains a Resource Manager (RM) and a mobility manager (MM). Gateway routers act as the interface between the CCN and the external IP network.

2. **Support for Seamless Terminal Mobility:** In order to provide wireless services anytime and anywhere, seamless terminal mobility must be supported in 4G networks. Mobile users will be able to roam across geographic boundaries of wireless networks through mobility management. There are two main issues in mobility management: location management and handover management. More details on mobility management is given in Section 2.2 as our research project mainly focuses on this area.

3. **Quality of Service:** Supporting multimedia applications with different QoS requirements in the presence of diversified wireless access technologies is another challenging issue for 4G wireless networks [BK10]. In such networks, depending on the bandwidth, mobility and application requirements, users will be able to switch among the different access technologies in a seamless manner. Efficient radio resource management and Call Admission Control (CAC) strategies will be key components in such heterogeneous wireless systems supporting various types of applications with different QoS requirements [NH05].

   Current QoS designs are usually made with a particular wireless system in mind [HY03]. For example, 3GPP has proposed a comprehensive QoS architecture for UMTS. It realized QoS in UMTS via the UMTS bearer service and its underlying bearer services [v.503]. However, providing QoS only in UMTS can not guarantee end-to-end QoS in 4G because systems that are non-UMTS are involved. To address this problem, internetworking with most common QoS architectures is studied in 3GPP.

4. **Security:** In a 4G open environment, various service providers and network operators are expected to share the core telecommunication infrastructure via end-user devices and open interfaces. The problem arises with such openness is much higher risk on security issues comparing to the traditional closed environment (e.g. public switched telephone networks) [PP07]. Hence, guaranteeing high level of security becomes another important issue to be tackled in the successful deployment of 4G networks.

   Existing security schemes for wireless systems in 2G and 3G are inadequate for 4G networks [AMMC02]. The key concern in security designs for 4G networks is flexibility. As mentioned in [HY03], "As existing security schemes are mainly designed for specific services such as voice service, they may not be applicable to 4G environments that will consist of many heterogeneous systems.". To design flexible security systems, some researchers are working on reconfigurable security mechanisms.

Mobile Terminals

In order to adapt to the larger variety of wireless networks, services and requirements of users in 4G systems, intelligent MTs are essential.

One important aspect to be conquered is the need of multiple antenna techniques. User terminals with multiple antennas need to be adopted to make full use of various wireless access technologies, and thus to achieve the goals of 4G systems such as high data rate, high reliability and long range communications. The multiple antenna technology helps 4G achieve those goals.

In [Tel99], Teletar demonstrated that using multiple antennas at both transmitter and receiver can dramatically increase channel capacity while the total transmit power is held constant. In [STT+02] the authors described a Multiple-Input Multiple-Output (MIMO) Orthogonal Frequency-Division Multiplexing (OFDM) wireless communication system which employs two transmit antennas and three receive antennas at the base station, and one transmit antenna and three receive antennas at the user terminal. Lab test results and field test results were also obtained. An overview of the multiplexing and scheduling techniques proposed in the context of multi-user MIMO-based wireless networks was provided in [AH05].

Besides multiple antenna techniques, other design problems such as limitations in device size, cost, power consumption, and backward compatibilities to systems are also to be tackled [HY03].

Services

4G systems will provide users with a wide variety of new services, and one customer may subscribe to many services from multiple service providers. Operators need to design new business architecture, accounting processes and accounting data maintenance for these services. Equalization on different charging schemes is also needed as different billing schemes may be used for different types of services (e.g., charging can be based on data, time or information). In order to build a structural billing system for 4G networks, several frameworks have already been studied [GV03, KK04].

1.2 Media Independent Handover Architecture

As mentioned in Section 1.1, one of the key challenges in 4G is the handover management in heterogeneous networks. In this section, the new specification for Media Independent Handover (MIH) services, IEEE 802.21 [DLOBS+08] is introduced.

The decision to initiate handovers in 3G networks has traditionally been based on the channel quality measured from the received signal strength and the availability of resources in the new cell. These are done periodically so that degradations in signal strength below a prescribed threshold can be detected and a handover to another radio channel or cell can be initiated. However, such traditional handover decisions use the signal strength as the only criteria. Also, traditional handovers do not allow users' selection of networks, and assume that there is only one choice of access technology. Traditional handover protocols are developed for homogeneous systems that rely on a common signaling protocol, routing technique and mobility management standard. In a heterogeneous environment, handover decisions could also be initiated for other reasons including bandwidth and security requirements. User choice is desirable, and mobile nodes and network routers must be able to inter-operate with different networks and protocols.

In order to enable seamless inter-system handovers, IEEE is currently working on a new specification for MIH services, IEEE 802.21. The aim of this specification is to improve user experience of MTs by enabling handovers between heterogeneous technologies while maintaining session continuity. IEEE 802.21 provides a framework that defines the interface between network layers, without having to deal with specifics of the technology implemented in any particular network layer. The general architecture of 802.21 is shown in Figure 1.2.

In this architecture, MIH Function (MIHF) acts as an intermediate layer between the upper and lower layers, allowing the exchange of information and commands between different devices which are involved in making handover decisions and executing handovers. Each node has a set of MIHF users, typically mobility management protocols, that use the MIHF functionality to control and gain handover-related information. Communications between MIHF and other functional entities, such as MIHF users and lower layers, are based on a number of defined service primitives which are grouped in Service Access Points (SAPs). MIHF provides three types of services:

1. Media Independent Event Service (MIES),
2. Media Independent Command Service (MICS), and

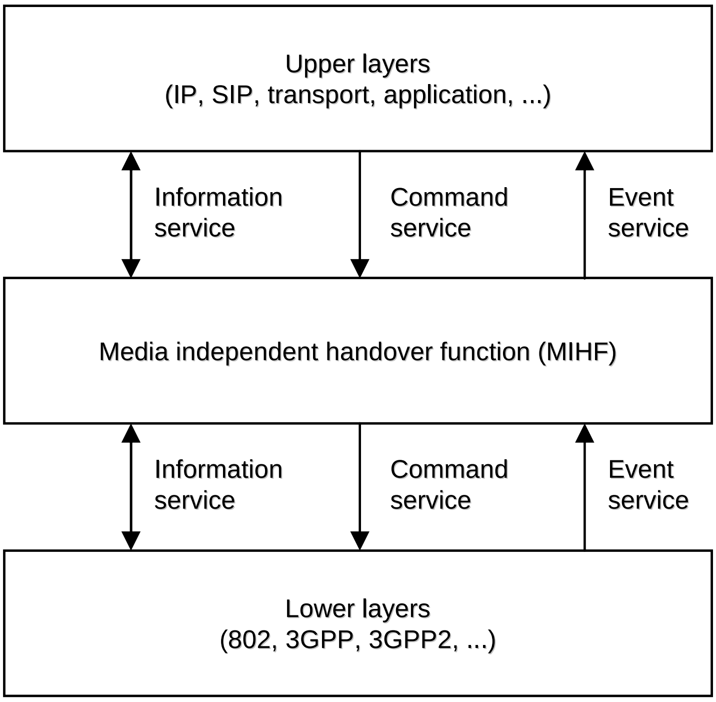

Figure 1.2: General architecture of IEEE 802.21. Three types of services are provided: event service, command service and information service.

3. Media Independent Information Service (MIIS).

MIES is responsible for detecting events and reporting them from both local and remote interfaces. This type of service is provided from lower layers to upper layers as depicted in Figure 1.2. Link deterioration and link unavailability are examples of such events that are reported to higher layers. On the other hand, the MICS defines commands for higher layers to control the lower layers regarding handovers. Commands follow a top-down direction as opposed to events. Typical commands are the configuration of network devices and the scanning of available networks. Less frequently used but equally important, MIIS which provides the mechanism for retrieving information and assisting the handover decision is also included in the set of service types. Such information can be static link layer parameters, like channel information, or the Medium Access Control (MAC) address of the Access Point (AP).

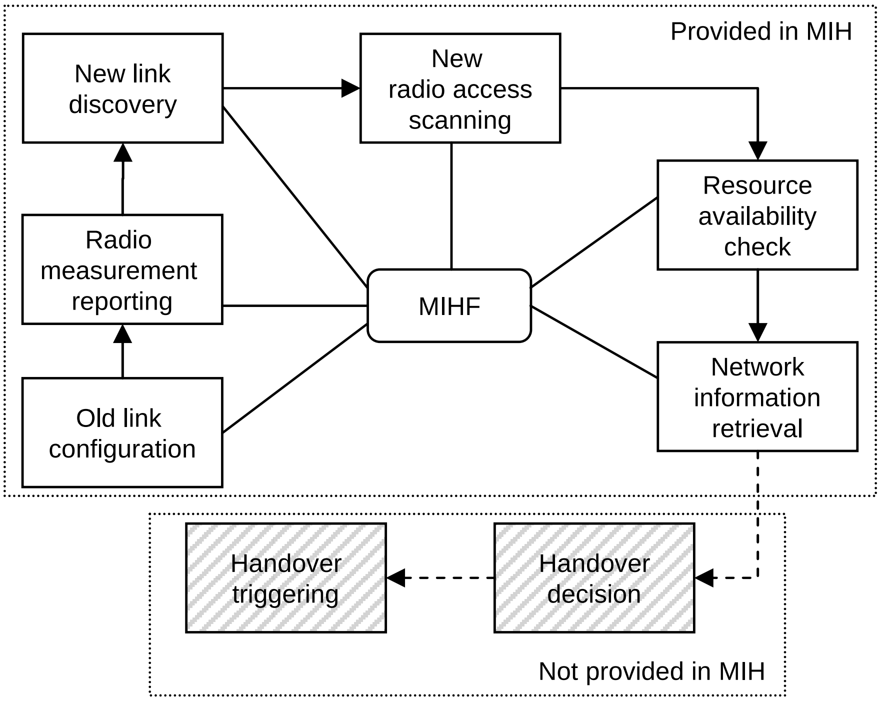

Figure 1.3: Handover in IEEE 802.21. The handover decision algorithm and handover execution are yet to be included in 802.21.

IEEE 802.21 offers handover procedures including old link configuration, radio measurement reporting, new link discovery, new radio access scanning, resource availability check and network information retrieval. However, the implementation of the handover decision algorithm and handover execution are not included in IEEE 802.21, as shown in Figure 1.3. In the 4G heterogeneous network environment, a handover management technique must choose the appropriate time to initiate the handover and the most suitable access network to handover to. The VHD process becomes especially important because it determines when and where to handover in a heterogeneous environment.

In this research project, a VHD scheme is developed to provide seamless mobility and better service quality for 4G users.

1.3 Optimization of Vertical Handover Decision (VHD) Processes

1.3.1 Thesis Objectives and Scope

A variety of VHD algorithms have been proposed to trigger handover at the optimal time to the optimal network based on a variety of network parameters. A detailed survey of these proposed algorithms can be found in the next chapter of this thesis.

These VHD algorithms either lack a comprehensive consideration of various network parameters or the studies reporting these algorithms lack enough detail for implementation. Besides that, there are two more problems with the existing VHD algorithms. The first one is that these algorithms tend to trigger handovers to low-cost and high-throughput WLANs whenever their coverage is available. In situations where the MT travels through an area close to the coverage boundary of a WLAN at speeds above a certain threshold, handovers to the WLAN will lead to network resource wastage as well as to the degradation of the MTs' battery life [CS05]. Furthermore, if the handover process has not been completed before the MT leaves the WLAN coverage area, a connection breakdown occurs. The second problem is that when the signal from the serving PoA is deteriorating and a handover is needed, the existing VHD algorithms for determining the handover triggering time are not able to dynamically adapt to user mobility and network parameters.

The research project presented in this thesis provides an optimized VHD scheme, which involves minimum number of handover failures, unnecessary handovers and connection breakdowns whilst maintaining a maximum user satisfaction. Such a scheme ensures the maximum connection time with a preferred access network with the minimum chance of service interruption. The scheme involves several VHD algorithms and chooses an algorithm intelligently based on conditions and user preferences.

1.3.2 Thesis Contributions

The VHD scheme presented in this thesis can be implemented in the IEEE 802.21 framework. The scheme consists of three VHD modules: Handover Necessity Estimation (HNE), Handover Target Selection (HTS), and Handover Triggering Condition Estimation (HTCE), as depicted in Figure 1.4. These three modules are described below:

1. **Handover necessity estimation (HNE):** A method which estimates the necessity of a handover is proposed. HNE includes two VHD algorithms. The first algorithm predicts the user's traveling time within a network coverage area, and the averaged Received Signal Strength (RSS) samples and the MT's velocity information are used in the traveling time prediction in a mathematical model. The second algorithm calculates a time threshold based on various network parameters and the handover failure or unnecessary handover probability information. The expression of handover failure or unnecessary handover probability is generated by developing a mathematical model which assumes uniform distribution of entry and exit points of a network coverage area. The predicted traveling time is compared against the time threshold and a handover is necessary only if the traveling time is longer than the threshold. This method leads to a reduction of handover failures of up to 80% and unnecessary handovers of up to 70%.

2. **Handover Target Selection (HTS):** A handover target selection method is

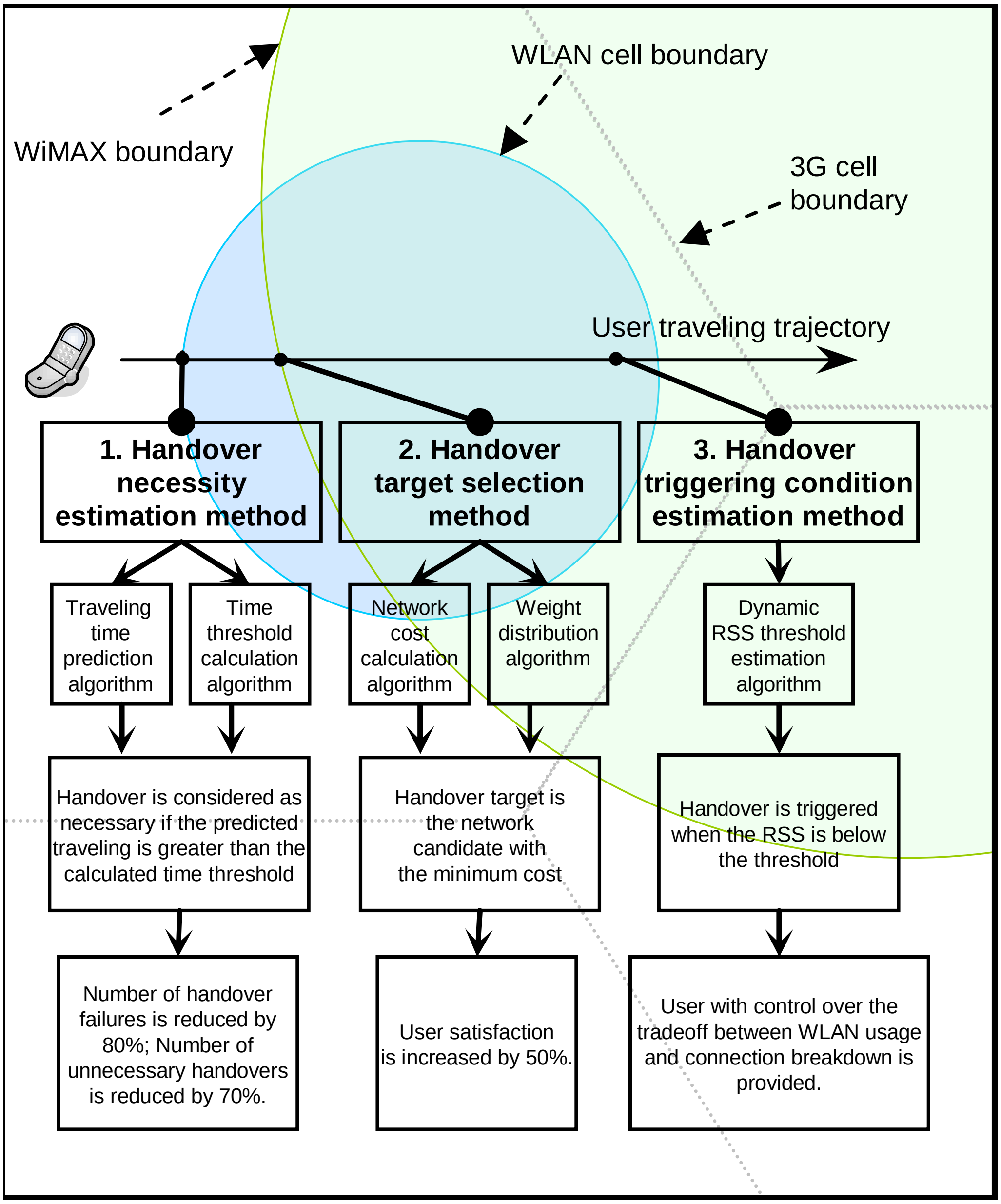

Figure 1.4: Contributions of the project. A scenario in which a user travels through an area with overlapping coverage of a 3G network, WLAN cell and WiMAX is used to explain the three main components of the project and their contributions.

presented to choose the best candidate network among all available networks. HTS adopts a cost function to calculate the cost of each candidate network and the handover target selected is the network with the minimum cost. The cost function involves various network parameters such as available bandwidth, power consumption, monetary cost and security level. A weight distribution unit is also included to assign different weights to the parameters based on user preferences and mobile status. This method increases the user satisfaction up to 50%, as measured using the method described in [HNH06].

3. **Handover Triggering Condition Estimation (HTCE):** A handover triggering condition estimation method is presented which helps to find the appropriate handover triggering condition when the RSS of the current serving network is fading. HTCE offers flexibility to either maximize WLAN usage or exert tight control on handover breakdowns to suit user's application requirements. The algorithm takes a parameter called "connection breakdown tolerance" as input, and estimates handover triggering time through statistical analysis of the probability of connection breakdown based on the MT speed and expected handover delay. HTCE is able to keep the connection breakdown probability below desirable limits, and provides the user with control over the tradeoff between connection breakdown probability and WLAN usage.

1.4 Thesis Outline

The structure of the thesis is as follows: A critical review of VHD algorithms reported in the research literature is presented in Chapter 2. In Chapter 3, overall framework of the proposed handover process optimization scheme is provided. Then, details of the three major components of the scheme, Handover Necessity Estimation (HNE), Handover Target Selection (HTS) and Handover Triggering Condition Estimation (HTCE) are explained in Chapters 4, 5 and 6, respectively. Theoretical and simulation based experimental results of the system are presented and discussed in Chapter 7. Conclusions of the research project, and suggestions for future research directions are given in Chapter 8.

Chapter 2

VHD Algorithms: State-of-the-Art

2.1 Introduction

Efficient VHD algorithms need to be designed to provide the required QoS to a wide range of applications while allowing seamless roaming among a multitude of access network technologies. In this chapter, a comprehensive survey of the VHD algorithms designed to satisfy these requirements is presented. To offer a systematic comparison, the algorithms are categorized into four groups based on the main handover decision criterion used. Also, to evaluate tradeoffs between their complexity of implementation and efficiency, three representative VHD algorithms in each group are discussed.

The rest of the chapter is organized as follows. Firstly, in Sections 2.2 and 2.3, background information on mobility management and an overview of VHD algorithms are provided. In Section 2.4, representative algorithms in four VHD groups are discussed. Then, comparisons among these four groups, and between the existing VHD algorithms and VHD algorithms proposed in this research project are presented in Section 2.5. In the last section, some conclusions are included.

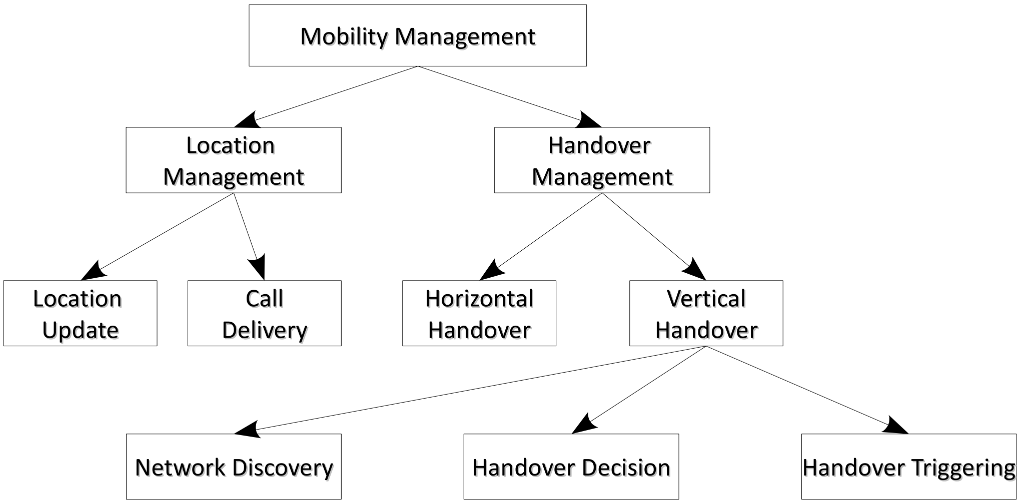

Figure 2.1: Mobility management in a heterogeneous network environment.

2.2 Handover Management in 4G Networks

As one of the most important challenges in 4G, mobility management is discussed in detail in this section. The hierarchy of mobility management in a heterogeneous network environment is depicted in Figure 2.1. There are two main areas for mobility management: location management and handover management [AMH+99].

Location management allows the network to discover the current Point of Attachment (PoA) of the mobile for call delivery [AMH+99]. It involves two stages, location update and call delivery. Location update or registration enables the network to authenticate the user and update the location of the mobile. In this stage, the MT periodically notifies the network of its new access point, allowing the network to authenticate the user and revise the user's location profile. This allows the network to keep track of the MT. Call delivery is responsible for database queries and terminal paging. In this stage, the network is queried for the user location profile and the current position of the MT is found.

Current techniques for location management involve database architecture design and the transmission of signaling messages between various components of a signaling network. Since location management deals with database and signaling issues, many of the issues are not protocol dependent and can be applied to various networks [HHK06].

Some recent work has reported on the location management in a multi-system environment. It has been shown in [AW02] that an integrated location management strategy can significantly outperform an independent operation of each sub-system's location management algorithm. In [GKCA08] the authors summarized various existing mobility management solutions and highlighted the view of mobility management issues in heterogeneous multi-hop wireless networks. In [MRD08], an integrated information-theoretic location management framework, which allows each individual sub-system to operate fairly independently, was developed for a multi-system environment.

In the 4G wireless environment, a mobile user is able to continue using the mobile device while moving from one point of attachment to another. Such process is called a handover, by which a mobile terminal keeps its connection active when it migrates from the coverage of one network access point to another [NHH06a]. In this section, some background information on handovers is provided.

Handover is the process of maintaining a user's active sessions when a mobile terminal changes its connection point to the access network (called "point of attachment"), for example, a base station or an access point [AMH+99]. Depending on the access network that each point of attachment belongs to, the handover can be either horizontal or vertical [NHH06b]. A horizontal handover takes place between points of attachment supporting the same network technology, for example, between two neighboring base stations of a cellular network. On the other hand, a vertical handover occurs between points of attachment supporting different network technologies, for example, between an IEEE 802.11 access point and a cellular network base station.

A handover process can be split into three stages: handover decision, radio link transfer and channel assignment [AMH+99]. Handover decision involves the decision to which point of attachment to execute a handover and its timing. Radio link transfer is the task of forming links to the new point of attachment, and channel assignment deals with the allocation of resources.

2.2.1 Classification of Handovers

Based on different factors used in the handover decision process, handovers can be classified in various ways. Some of the popular classifications are discussed below.

Horizontal and Vertical Handover - Depending on the network types involved, handovers can be classified as either horizontal or vertical [ZM06]. A horizontal handover or intra-system handover takes place between PoA supporting the same network technology, e.g., two geographically neighboring BSs of a 3G cellular network. On the other side, a vertical handover or inter-system handover occurs between PoA supporting different network technologies, e.g., an IEEE 802.11 AP and a 3G BS. An example of horizontal and vertical handovers is illustrated in Figure 2.2, where a horizontal handover happens between two cellular BSs and a vertical handover takes place between an AP of a WLAN and a BS of a cellular BS.

Vertical handovers are implemented across heterogeneous cells of access systems,

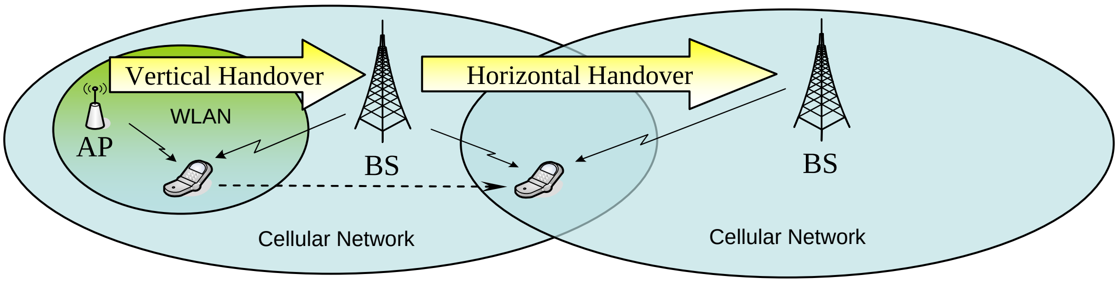

Figure 2.2: An example of horizontal and vertical handovers in heterogeneous wireless networks.

which differ in several aspects such as bandwidth, data rate, frequency of operation, etc. The different characteristics of the networks involved make the implementation of vertical handovers more challenging as compared to horizontal handovers [SZ06]. The terms horizontal and vertical follow from the overlay network structure that has networks with increasing cell sizes at higher levels in the hierarchy. Vertical handovers are generally of two types namely, upward and downward handovers.

An Upward vertical handover is a handover to a wireless overlay with a larger cell size and generally lower bandwidth per unit area. So, an upward vertical handover makes a mobile device disconnect from a network providing faster but smaller coverage (example WLAN) to a new network providing slower but broader coverage. A downward vertical handover is a handover to a wireless overlay with a smaller cell size, and generally higher bandwidth per unit area. A mobile device performing a downward vertical handover disconnects from a cell providing broader coverage to one providing limited coverage but higher access speed [NHH06a].

Hard and Soft Handover - Based on the number of connections involved, handovers are classified as hard handovers and soft handovers. A handover is hard if the MT can be associated with only one access point at a time. A soft handover occurs if the MT can communicate with more than one access points during the handover. For example, if the MT is equipped with multiple network interfaces, it can simultaneously connect to multiple APs in different networks during soft handover [SZ06]. Soft handover may also be referred to as make-before-break handover in which the mobile node's connection may be created at the target BS before the old BS connection is released. On the other hand, in the case of hard handover or break-before-make handover, the new connection may be set up after the old connection has been torn down.

Mobile-controlled, Network-controlled and Mobile-assisted Handover - Under network-controlled handover, the network makes the decision for handover, while under mobile-controlled handover, the mobile node must make the handover decision on its own. Under mobile-assisted handover, the decision to handover is made by the mobile node in cooperation with the network.

Other Handover Classifications - Besides the ways of classifying handovers stated above, there are other methods to categorize handovers [NHH06a]. Based on the factor of frequencies engaged, there are intra-frequency handovers and inter-frequency handovers. For different administrative domains engaged, there are intra-administrative handovers and inter-administrative handovers. Depending on the necessity of handovers, handovers are divided into obligatory handovers and voluntary handovers. The last factor is user control allowance, and there are proactive and passive handovers.

2.2.2 Vertical Handover Process

Three stages are involved in a vertical handover process: network discovery, handover decision, and handover triggering [AW02].

Network Discovery - This is the process where a MT searches for reachable wireless networks. A MT with multiple interfaces must activate the interfaces to receive service advertisements, which are broadcasted by different wireless technologies. The MT will know a wireless network is reachable if its service advertisements can be heard. The simplest way to discover reachable wireless networks is to always keep all interfaces on. However, keeping an interface active all the time consumes the battery power even without receiving or sending any packets. Therefore, to avoid keeping the idle interfaces always on is critical. Also the discovery time should be low so that the MT can benefit faster from the new wireless network.

The power efficiency and the system discovery time are the most critical considerations for system discovery methods' performance. The interface may be activated periodically to receive service advertisements. The activating frequency directly affects the system discovery time. The MT that activates the interfaces with high frequency may discover the reachable network quickly but its battery may run out very soon. The MT that activates the interface with low frequency may increase the power efficiency, but it may discover the reachable wireless networks slowly. There exists a tradeoff between the power efficiency and the system discovery time [CLH04].

Handover Decision - Handover decision is the ability to decide when to perform the handover and to which access network to handover. A decision for vertical handover may depend on several issues relating to the network to which the mobile node is already connected and to the one that it is going to handover. For example, the decision to perform mobile-controlled handovers may be made by a vertical handover agent, sitting in the mobile device based on policies such as network bandwidth, load, coverage, cost, security, QoS, or even user preferences [ZM04].

More details on vertical handover decisions are discussed in the Section 2.3.

Handover Triggering - Handover triggering requires the actual transfer of data packets to a new wireless link in order to re-route a mobile user's connection path to the new PoA. It requires the network to transfer routing information about the mobile user to the new (or target) access router for the proper forwarding of packets. Since 4G heterogeneous networks will operate in an environment of multiple standards and networks, transfer of packets to a new wireless link will also involve transfer of additional contextual information in order to enable the mobile node to move through different networks, while maintaining its data flows. The desired goal of transferring the context of a mobile node to the new network is to minimize the delay in re-establishing the mobile node's traffic flows. However, if the context transfer delay is so large as to have the same effect of the complete re-establishment, or large enough to increase the overall handover call dropping rate, the advantages of context transfer are lost. Thus, a mechanism to allow for inter-network and/or inter-service-provider agreements to support fast inter-system handovers, while avoiding an unreasonable amount of inter-network signaling exchanges to validate or institute the adjustment in services, is presently a crucial research problem [ZM04].

VHD algorithms help mobile terminals to choose the best network to connect to among all the available candidates. Here, the focus is only on the research efforts and recent developments on improving the efficiency of VHD process. In contrast to horizontal handover decision algorithms which mainly consider RSS as the only decision criterion, for VHD algorithms, criteria such as cost of services, power consumption and velocity of the mobile terminal may need to be taken into consideration to maximize user satisfaction [NHH06b].

From the following sections, the focus is on VHD algorithms.

2.3 Overview of VHD Algorithms

A number of studies published earlier have surveyed VHD algorithms [MZ04, ZL05, SNW06]. In the earliest one [MZ04], a tutorial on the design and performance issues of VHD policies is presented along with the analysis and comparison of several VHD algorithms. However, the focus of this study was quite narrow and only covered cost function and RSS based VHD algorithms. In a later study [ZL05], the authors presented a framework to compare the performance of different vertical handover algorithms on system resource utilization and QoS perceived by users, but only included the evaluation of two VHD algorithms. A subsequent survey's focus [SNW06] was on various mathematical models used in vertical handover decisions. In the rest of this chapter, the earlier studies are updated by incorporating recently published algorithms.

In this section, the existing VHD algorithms are firstly categorized into four groups based on the main handover decision criterion used.

2.3.1 VHD Criteria

Several parameters as shown in Figure 2.3 have been proposed in the research literature for use in the VHD algorithms. We briefly explain each of them below.

Received signal strength (RSS) is the most widely used criterion because it is easy to measure and is directly relevant to the service quality. There is a close relationship between the RSS readings and the distance between the mobile terminal and its point of attachment. The majority of existing horizontal handover algorithms use RSS as the main decision criterion, and RSS is an important criterion for VHD algorithms.

Network connection time refers to the duration that a mobile terminal remains

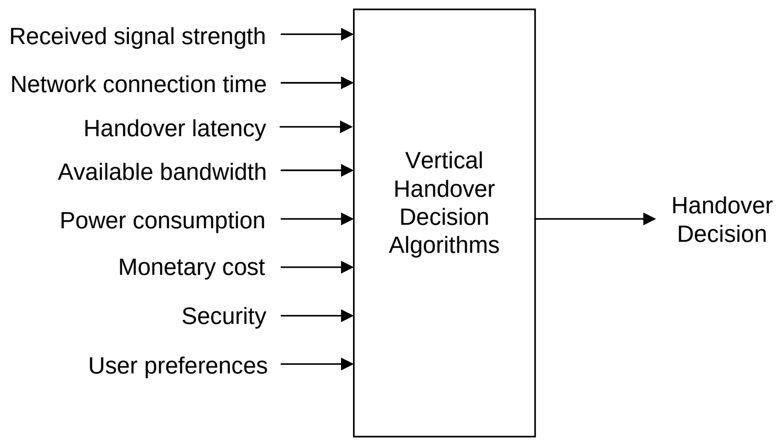

Figure 2.3: Parameters used for making VHD decisions.

connected to a particular access network. Determining the network connection time is very important for choosing the right moment to trigger a handover so that the service quality could be maintained at a satisfactory level. For example, a handover done too early from a WLAN to a cellular network would waste network resources or being too late would result in a handover failure. Determining the network connection time is also important for reducing the number of superfluous handovers, as handing over to a target network with potentially short connection time should be discouraged.

The network connection time is related to a mobile terminal's location and velocity. Both the distance from the mobile terminal to its point of attachment and the velocity of the mobile terminal affect the RSS at the mobile terminal. The variation of the RSS then determines the time in which the mobile terminal stays connected to a particular network. Network connection time is especially important for VHD algorithms because heterogeneous networks usually have different sizes of network coverage.

Handover latency is defined for a MT as the time that elapses between the last packets received via the old access router and the arrival of the first packet along the new access router after a handover. Handover latency can be considerably different between various technologies and this has a major impact on interactive applications.

Available bandwidth is a measure of available data communication resources expressed in bit/s. It is a good indicator of the traffic conditions in the access network.

Power consumption becomes a critical issue especially if a mobile terminal's battery is low. In such situations, it would be preferable to hand over to a point of attachment which would help extending valuable battery life [NHH06b].

Monetary cost: For different networks, there would be different charging policies, therefore, in some situations the cost of a network service should be taken into consideration in making handover decisions.

Security: For some applications, confidentiality or integrity of the transmitted data can be critical. For this reason, a network with higher security level may be chosen over another one which would provide lower level of data security.

User preferences: A user's personal preference towards an access network could lead to the selection of one type of network over the other candidates.

RSS and network connection time based decision criteria are widely used in both horizontal and vertical handover decisions. Others are mainly seen in VHD schemes only.

2.3.2 Classification of VHD Algorithms

There are various ways to classify VHD algorithms [KKP08, LWL08]. In this dissertation, VHD algorithms are divided into four groups based on the handover decision criteria used and the methods used to process them.

RSS based algorithms: RSS is used as the main handover decision criterion in this group. Various strategies have been developed to compare the RSS of the current point of attachment with that of the candidate point of attachment [ZLS06, MA06, YMŞ08]. In [Pol96] RSS based horizontal handover decision strategies are classified into the following six subcategories: relative RSS, relative RSS with threshold, relative RSS with hysteresis, relative RSS with hysteresis and threshold, and prediction techniques. For VHD, relative RSS is not applicable, since the RSS from different types of networks can not be compared directly due to the disparity of the technologies involved. For example, separate thresholds for each network. Furthermore, other network parameters such as bandwidth are usually combined with RSS in the VHD process.

Bandwidth based algorithms: Available bandwidth for a mobile terminal is the main criterion in this group [LCCS05, YGQD07, CCHL07]. In some algorithms, both bandwidth and RSS information are used in the decision process [ZLS06, GGZZ04]. Depending on whether RSS or bandwidth is the main criterion considered in the algorithm, in this survey, the method is classified either as RSS based or bandwidth based.

Cost function based algorithms: This class of algorithms combine metrics such as monetary cost, security, bandwidth and power consumption in a cost function, and the handover decision is made by comparing the result of this function for the candidate networks [ZM04, HNH06, TPS08]. Different weights are assigned to different input metrics depending on the network conditions and user preferences.

Combination algorithms: These VHD algorithms attempt to use a richer set of inputs than the others for making handover decisions. When a large number of inputs are used, it is usually very difficult or impossible to develop analytical formulations of handover decision processes. Due to this reason, researchers apply machine learning techniques to formulate the processes. Our literature survey reveals that fuzzy logic and artificial neural networks based techniques [Zha04, PKH+00] are popular choices. Fuzzy logic systems allow human experts' qualitative thinking to be encoded as algorithms to improve the overall efficiency. Examples of applying this approach into VHD can be found in [XJH07, CSH+01, Zha04, HO06, LTD06]. If there is a comprehensive set of input-desired output patterns available, artificial neural networks can be trained to create handover decision algorithms [GZX05, NGAM07, PKH+00]. It is also possible to create adaptive versions of these algorithms which, through continuous and real time learning processes, the systems can monitor their performance and modify their own structure to create highly effective handover decision algorithms.

2.3.3 Performance Evaluation Metrics for VHD Algorithms

VHD algorithms can be quantitatively compared under various usage scenarios by measuring the mean and maximum handover delays, the number of handovers, the number of failed handovers due to incorrect decisions, and the overall throughput of a session maintained over a typical mobility pattern. These metrics are further explained below:

Handover delay: Handover delay is the duration between the initiation and completion of the handover process, and is related to the complexity of the VHD process. Reduction of the handover delay is especially important for delay sensitive voice or multimedia applications.

Number of handovers: Reducing the number of handovers is usually preferred as frequent handovers would cause wastage of network resources. A handover is considered to be superfluous when a handover back to the original point of attachment is needed within a certain time duration [CRMRS99, YMŞ08], and the number of such handovers should be minimized.

Handover failure probability: A handover failure occurs when the handover is initiated but the target network does not have sufficient resources to complete it, or when the mobile terminal moves out of the coverage of the target network before the process is finalized. In the former case, the handover failure probability is related to the channel availability of the target network [XT04], while in the latter case it is related to the mobility of the user [Bar04].

Throughput: The throughput refers to the data rate delivered to the mobile terminals on the network. Handover to a network candidate with higher throughput is usually desirable.

2.4 Representative Vertical Handover Decision Algorithms

In this section, a representative set of VHD algorithms is discussed. These algorithms are selected because they make a good representation of their VHD groups. Their operational fundamentals are summarized along with their comparative advantages and disadvantages. These algorithms are assigned into one

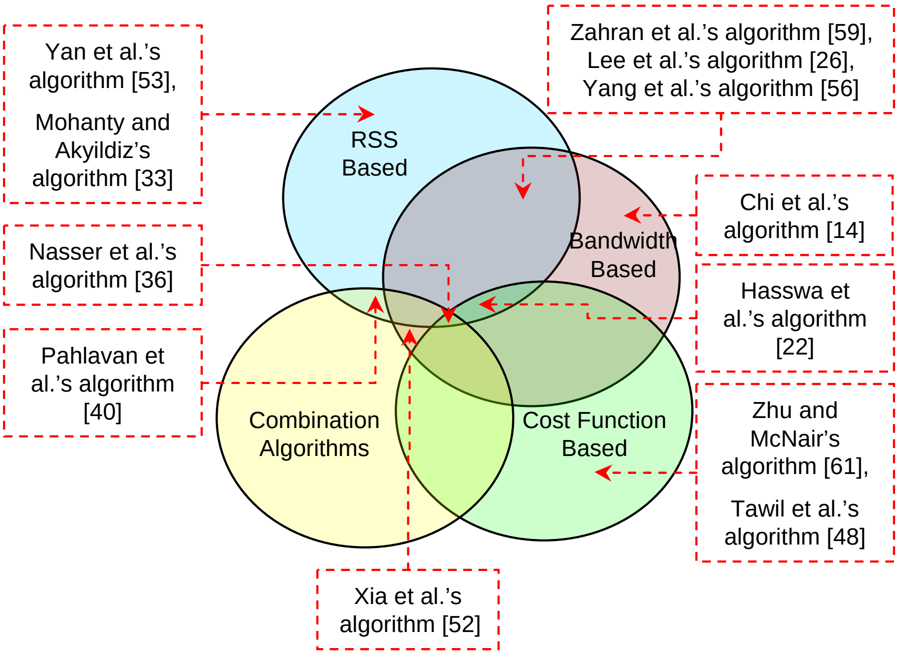

Figure 2.4: The four categories of VHD algorithms and twelve selected representative schemes.

of the four categories described in Section 2.3.2 (as shown in Figure 2.4). Some of the algorithms use more than one VHD criteria, and in such cases, only the main criterion used for classification is considered.

2.4.1 RSS Based VHD Algorithms

RSS based VHD algorithms compare the RSS of the current point of attachment against the others to make handover decisions. Because of the simplicity of the hardware required for RSS measurements, not surprisingly, a large number of studies have been conducted in this area [PYK+03, LLGL06, ZLS06, MA06, YMŞ08, CC08]. Three representative RSS based VHD algorithms are described in the following sections.

An Adaptive Lifetime Based Handover Heuristic

Zahran et al. [ZLS06] proposed an algorithm for handovers between 3G networks and WLANs by combining the RSS measurements either with an estimated lifetime metric (expected duration after which the mobile terminal will not be able to maintain its connection with the WLAN) or the available bandwidth of the WLAN candidate. Their method is described through the following scenarios:

In the first scenario, when the mobile terminal moves away from the coverage area of a WLAN into a 3G cell, a handover to the 3G network is initiated. The handover is triggered under the conditions that (a) RSS average of the WLAN connection falls below a predefined threshold (MOT_{WLAN}), and (b) the estimated lifetime is less than or equal to the handover delay. The mobile terminal continuously calculates the RSS average using the moving average method

$$\overline{\text{RSS}}[k] = \frac{1}{W_{\text{av}}} \sum_{i=0}^{W_{\text{av}}-1} \overline{\text{RSS}}[k-i]. \tag{2.4.1}$$

Here $\overline{\text{RSS}}[k]$ is the calculated average of RSS at time instant k, and W_{av} is the window size of a slope estimator, a variable that changes with the velocity of the mobile terminal. Then, the lifetime metric $\text{EL}[k]$ is calculated by using $\overline{\text{RSS}}[k]$, the RSS change rate $S[k]$, and a parameter called Application Signal Strength Threshold (ASST) as follows

$$\text{EL}[k] = \frac{\overline{\text{RSS}}[k] - \text{ASST}}{S[k]}. \tag{2.4.2}$$

The RSS change rate $S[k]$ varies with the window size of the slope estimator and the RSS sampling interval. For details on calculating $S[k]$, please refer to equations (4), (5) and (6) in [ZLS06]. The ASST is an application dependent parameter

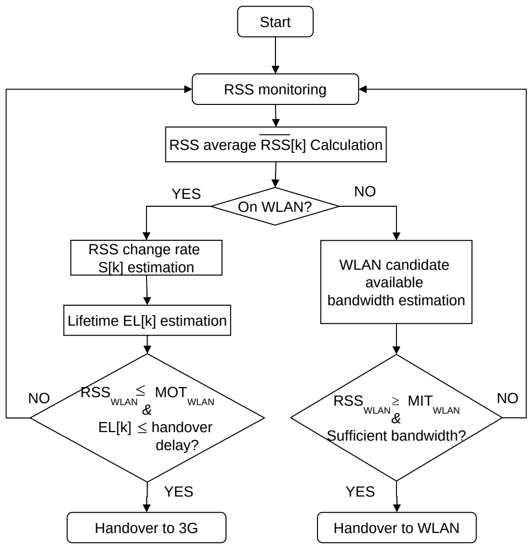

Figure 2.5: Zahran et al.'s VHD algorithm [ZLS06].

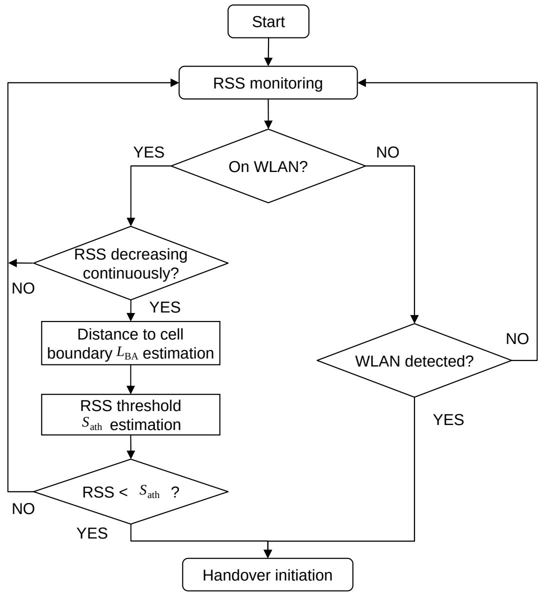

Figure 2.6: Mohanty and Akyildiz's VHD heuristic [MA06].

which represents a composite of the channel bit error rate, application error resilience and application QoS requirements. A lookup table for the optimal ASST values is provided in the paper.

In the second scenario, when the mobile terminal moves towards a WLAN cell, the handover to the WLAN is triggered if the average RSS measurements of the WLAN signal is larger than a threshold (MIT_{WLAN}) and the available bandwidth of the WLAN meets the bandwidth requirements of the application. The flowchart of Zahran et al.'s heuristic is depicted in Figure 2.5.

Benefits of Zahran et al.'s algorithm can be summarized as follows. First, by introducing the lifetime metric, the algorithm adapts to the application requirements and the user mobility, reducing the number of superfluous handovers significantly. Second, there is an improvement on the average throughput for the user because of the mobile terminal's ability to remain connected to the WLAN cell as long as possible. However, packet delays grow with an increase in the lifetime, due to the deterioration of the channel condition as the mobile terminal approaches the edge of the WLAN cell. This issue can be critical for delay sensitive applications and degrade their performance. To solve this problem, ASST is tuned according to various system parameters, including delay thresholds, mobile terminal velocities, handover signaling costs and packet delay penalties.

An RSS Threshold Based Dynamic Heuristic

Mohanty and Akyildiz [MA06] proposed a WLAN to 3G handover decision method based on comparison of the current RSS and a dynamic RSS threshold (S_{dth}) when a mobile terminal is connected to a WLAN access point. S_{dth} (in dBm) is calculated as

$$S_{\text{dth}} = \text{RSS}_{\text{min}} + 10\beta \log 10 \left(\frac{d}{d - L_{\text{BA}}} \right) + \epsilon \tag{2.4.3}$$

where RSS_{min} (in dBm) is the minimum level of the RSS required for the mobile terminal to communicate with an access point, β is the path loss coefficient, d is the side length of the WLAN cell (in meters, a WLAN cell is assumed to have a hexagonal shape in this study), L_{BA} is the shortest distance between the point at which handover is initiated and WLAN boundary, and ϵ (in dB) is a zero-mean Gaussian random variable with a standard deviation that represents the statistical variation in RSS caused by shadowing. The distance L_{BA} changes with the tolerable handover failure probability p_f, the velocity of the mobile terminal v, and the WLAN to 3G handover delay τ, and calculated as

$$L_{\text{BA}} = \left[\tau^2 v^2 + d^2 \left(p_f - 2 + 2\sqrt{1 - p_f} \right) \right]^{\frac{1}{2}}. \tag{2.4.4}$$

The use of a dynamic RSS threshold helps reducing the incidences of false handover initiation and keeping the handover failures below a limit. However, in this algorithm, the handover failure probability from 3G network to a WLAN cell is considered to be zero since the 3G network coverage is assumed to be available all the time, and thus according to the mechanism, a handover to a WLAN is always desirable whenever the mobile terminal enters the WLAN coverage. Yan et al. [YMŞ08] (discussed in the next session) point out in their study that this is not efficient when the mobile terminal's traveling time inside a WLAN cell is less than the handover delay, and in such cases a handover may result in wastage of network resources.

A Traveling Distance Prediction Based Heuristic

To eliminate unnecessary handovers in the method presented in Section 2.4.1, Yan et al. [YMŞ08, YŞM08, YŞN] developed a VHD algorithm that takes into consideration the time the mobile terminal is expected to spend within a WLAN cell. The method relies on the estimation of WLAN traveling time (i.e. time that the mobile terminal is expected to spend within the WLAN cell) and the calculation of a time threshold (T_{WLAN}). A handover to a WLAN is triggered if the WLAN coverage is available and the estimated traveling time inside the WLAN cell is larger than the time threshold. The estimated traveling time (t_{WLAN}) is

$$t_{\text{WLAN}} = \frac{R^2 - l_{\text{OS}}^2 + v^2(t_s - t_{P_i})^2}{v^2(t_s - t_{\text{in}})} \tag{2.4.5}$$

where R is the radius of the WLAN cell, l_{OS} is the distance between the access point and where the mobile terminal takes an RSS sample, v is the velocity of the mobile terminal, and t_s and t_{P_i} are the times at which the RSS sample is taken and the mobile terminal enters the WLAN cell coverage, respectively. l_{OS} is estimated by using the RSS information and log-distance path loss model.

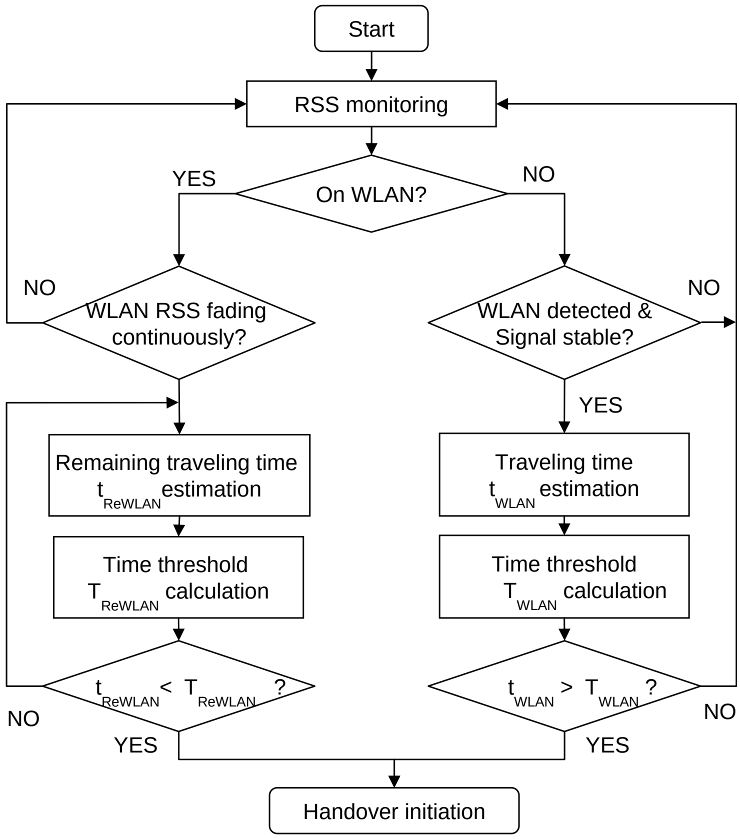

Figure 2.7: Yan et al.'s VHD heuristic [YMŞ08, YŞM08, YŞN].

The time threshold (T_{WLAN}) is calculated based on various network parameters as

$$T_{\text{WLAN}} = \frac{2R}{v} \sin\Big(\sin^{-1}\Big(\frac{v\tau}{2R}\Big) - \frac{\pi}{2}P\Big) \tag{2.4.6}$$

where τ_i is the handover delay from the cellular network to the WLAN, and P is the tolerable handover failure, unnecessary handover or connection breakdown probability. A handover to the cellular network is initiated if the WLAN RSS is continuously fading and the mobile terminal reaches a handover commencement boundary area which size is dynamic to the mobile terminal's speed. Figure 2.7 shows Yan et al.'s heuristic.

Heuristic	Applicable Area	Feature	Advantages	Disadvantages
Zahran et al.'s heuristic [59]	Between 3G and WLANs	The RSS is combined with an estimated lifetime or the available bandwidth to decide the handover time	• Adaptation to application requirements and user mobility • Improvement on the available bandwidth	• Long packet delay • Extra lookup table
Mohanty and Akyildiz's heuristic [33]	Between 3G and WLANs	A dynamic RSS threshold is calculated and compared with the current RSS to determine the handover time	• Reduction of the false handover initiation and handover failure probabilities	• Increased handover failure • Wastage of network resources
Yan et al.'s heuristic [53]	Between cellular networks and WLANs	A dynamic time threshold is calculated and compared with the predicted traveling time inside the WLAN to help with handover decisions	• Minimization of the handover failure, unnecessary handover and connection breakdown probabilities	• Extra handover delay

Table 2.1: A summary of RSS based VHD algorithms.

The main advantage of this heuristic is that it minimizes handover failures, unnecessary handovers and connection breakdowns. But the method relies on sampling and averaging RSS points, which introduces increased handover delay. The performance on the handover delay should be further discussed and balanced against the probability of unnecessary handovers.

A summary of the RSS based VHD heuristics is shown in Table 2.1.

2.4.2 Bandwidth Based VHD Algorithms

Bandwidth based VHD algorithms consider available bandwidth for a mobile terminal or traffic demand as the main criterion [NWDZ05, LCCS05, YGQD07, CCHL07]. In this section, three typical bandwidth based VHD algorithms are discussed in detail.

A QoS Based Heuristic

Lee et al. [LCCS05] devised a QoS based VHD algorithm which takes residual bandwidth and user service requirements into account in deciding whether to handover from a WLAN to Wireless Wide Area Network (WWAN) and vice versa.

When the mobile terminal is connected to a WLAN, the handover algorithm is initiated if the measured RSS is consistently below a threshold (RSS_{T1}). The algorithm also takes the state of the mobile terminal into consideration. If the mobile terminal is in the idle state, a handover to the preferred access network is performed, otherwise the handover decision is based upon the user application type. For delay-sensitive applications, a handover occurs only if the current serving WLAN is not able to provide enough bandwidth for the application while the WWAN is able to provide the necessary bandwidth. For delay-tolerant applications, a handover takes place if the WWAN provides higher bandwidth than the WLAN. An approximate value of the residual bandwidth of the WLAN is evaluated by the following formula:

$$\begin{aligned}\text{residual\_bandwidth} =& \, throughput \times (1 - \alpha \times channel\_utilization) \\ & \times (1 - packet\_loss\_rate) \end{aligned} \tag{2.4.7}$$

where $throughput$ is the throughput that can be shared among mobile terminals in the WLAN, $channel\_utilization$ is the percentage of time the access point senses the medium is busy using the carrier sense mechanism, α is a factor that reflects IEEE 802.11 MAC overhead (it is set to 1.25 in this paper), and $packet\_loss\_rate$ is the portion of transmitted medium access control (MAC) protocol data units (MPDUs) that require retransmission, or are discarded as undeliverable. The values of $channel\_utilization$ and $packet\_loss\_rate$ are obtained from the information in the beacon frame carrying the QoS basic service set (QBSS) load sent by an access point, as defined in the IEEE 802.11e [DStP03].

When the mobile terminal is connected to a WWAN, a similar process is carried

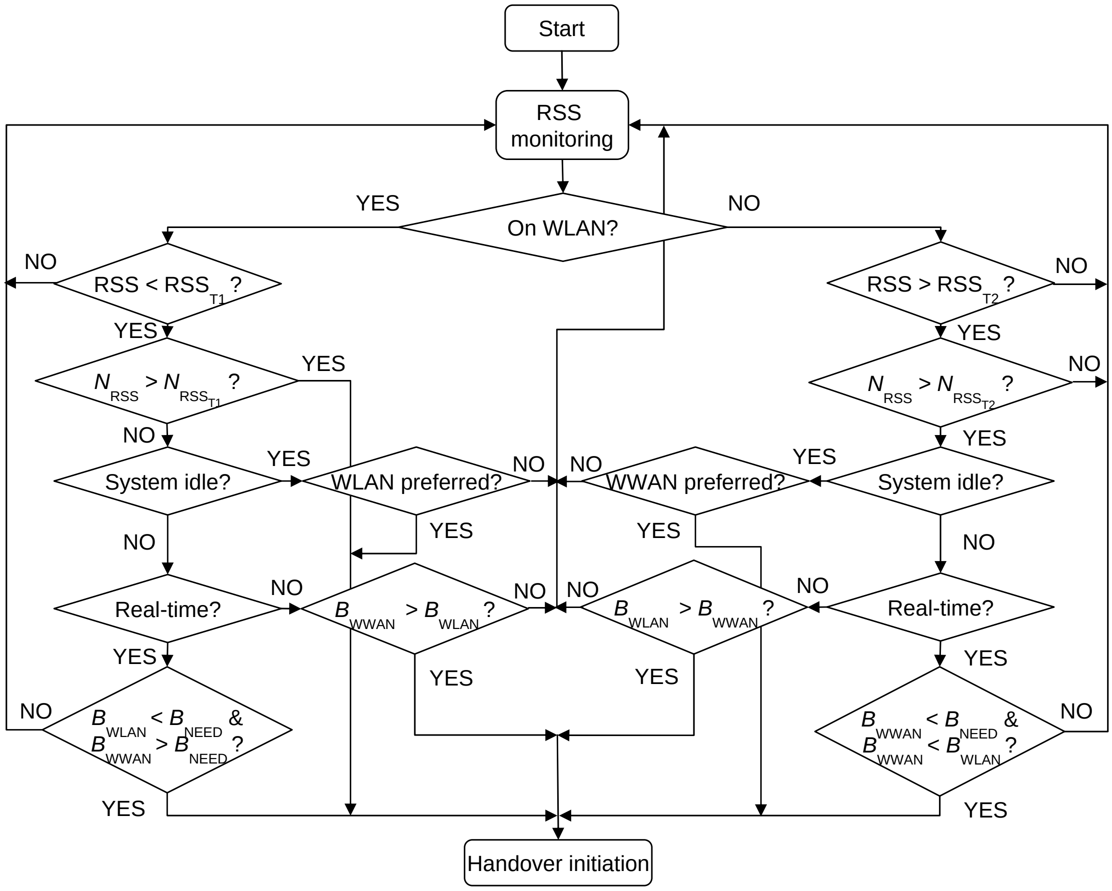

Figure 2.8: Lee et al.'s VHD heuristic [LCCS05].

out if consecutive beacons from the WLAN with RSS above a threshold (RSS_{T2}) are received. The flowchart of Lee et al.'s algorithm is depicted in Figure 2.8.

By considering the available bandwidth as the main VHD criterion, this heuristic is able to achieve high system throughput, and by taking application types into account, lower handover latency for delay-sensitive applications is achieved. However, acquiring the available bandwidth information in a cellular network for handover decisions is difficult [LCCS05]. Furthermore, in this method, a handover to the preferred network is performed when the mobile terminal is in the idle state. However, when the mobile terminal is staying in the preferred network for only a short period, the movement can result in high blocking rate for new applications.

A Signal to Interference and Noise Ratio (SINR) Based Heuristic

Yang et al. [YGQD07] presented a bandwidth based VHD method between WLANs and a Wideband Code Division Multiple Access (WCDMA) network using Signal to Interference and Noise Ratio (SINR). The SINR calculation of the WLAN signals is converted to an equivalent SINR to be compared with the SINR of the WCDMA channel

$$\gamma_{\text{AP}} = \Gamma_{\text{AP}} \left[(1 + \frac{\gamma_{\text{BS}}}{\Gamma_{\text{BS}}})^{\frac{W_{\text{BS}}}{W_{\text{AP}}}} - 1 \right] \tag{2.4.8}$$

where γ_{AP} and γ_{BS} are the SINR at the mobile terminal when associated with WLAN and WCDMA, respectively. Γ is the dB gap between the uncoded Quadrature Amplitude Modulation (QAM) and channel capacity, minus the coding gain, and Γ_{AP} equals to 3dB for WLAN and Γ_{BS} equals to 3dB for WLAN, as stated by the authors. W_{AP} and W_{BS} are the carrier bandwidth of WLAN and WCDMA links. A handover to the network with larger SINR is performed, as shown in the flowchart (Figure 2.9).

SINR based handovers can provide users with higher overall throughput than RSS based handovers since the available throughput is directly dependent on the SINR, and this algorithm results in a balanced load between the WLAN and the WCDMA networks. But such an algorithm may also introduce excessive handovers with the variation of the SINR causing the node to hand over back and forth between two networks, commonly referred to as ping-pong effect [Pol96].

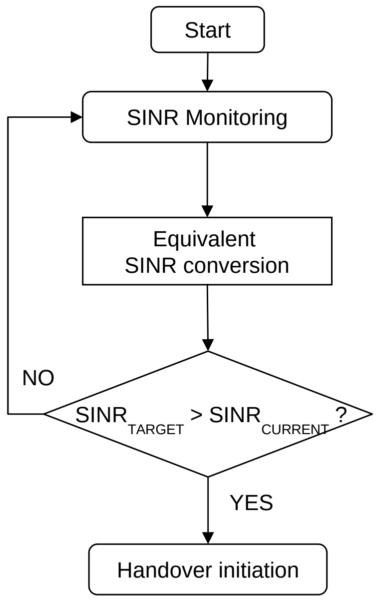

Figure 2.9: Yang et al.'s VHD heuristic [YGQD07].

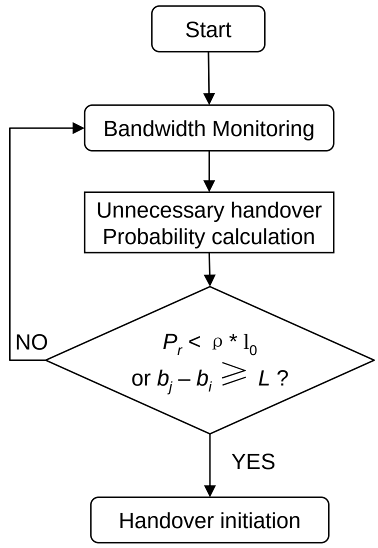

Figure 2.10: Chi et al.'s VHD heuristic [CCHL07].

A Wrong Decision Probability (WDP) Prediction Based Heuristic

In [CCHL07], Chi et al. proposed a VHD heuristic based on the Wrong Decision Probability (WDP) prediction. The WDP is calculated by combining the probability of unnecessary handovers and the missing handovers. Assume that there are two networks i and j with overlapping coverage, and b_i and b_j are their available bandwidth. An unnecessary handover occurs when the mobile terminal is in network i and decides to handover to j, but b_j is less than b_i after this decision. A missing handover occurs when the mobile terminal decides to stay connected to network i, but b_i is less than b_j after this decision.

A handover from network i to network j is initiated if $P_r < \rho \times l_0$ or $b_j - b_i \leq L$, where P_r is the unnecessary handover probability, ρ is the traffic load of network i, $l_0 = 0.001$, and L is a bandwidth threshold. The flowchart of this algorithm is shown in Figure 2.10.

The authors show that this algorithm is able to reduce the WDP and balance the

Heuristic	Applicable Area	Feature	Advantages	Disadvantages
Lee et al.'s heuristic [26]	Between WWANs and WLANs	The bandwidth is combined with the RSS, system status and application type to make handover decisions	• High system throughput • Low handover latency for real-time transmission	• Difficulty in acquiring available bandwidth information • Increased new application blocking rate
Yang et al.'s heuristic [56]	Between WCDMA and WLANs	The SINR values are compared to determine the handover decision	• High overall throughput • Balance of the network load between WLANs and WCDMA	• Excessive handovers • Ping-pong effect
Chi et al.'s heuristic [14]	Between any two wireless networks	Available bandwidth, network traffic and unnecessary handover probability are considered in the handover decision criteria	• Reduced unnecessary handover probability • Balance of the traffic load	• Increased connection breakdown probability without considering the RSS

Table 2.2: A summary of bandwidth based VHD algorithms.

traffic load, however, RSS is not considered. A handover to a target network with high bandwidth but weak received signal is not desirable as it may bring the connection breakdown.

A summary of the bandwidth based VHD heuristics is shown in Table 2.2.

2.4.3 Cost Function Based VHD Algorithms

The cost function based algorithms combine metrics in a cost function. Many studies have been done in this area [BI04, CSC$^+$04, ZM04, CS05, LHH06, HNH06, TPS08, SNLW08, LSK$^+$09]. In this section, three representative cost function based VHD algorithms are evaluated.

A Multi-service Based Heuristic

Zhu and McNair's [ZM04, ZM06] VHD algorithm relies on a cost function which calculates the "cost" of possible target networks. The algorithm prioritizes all the active applications, and then the cost of each possible target network for the

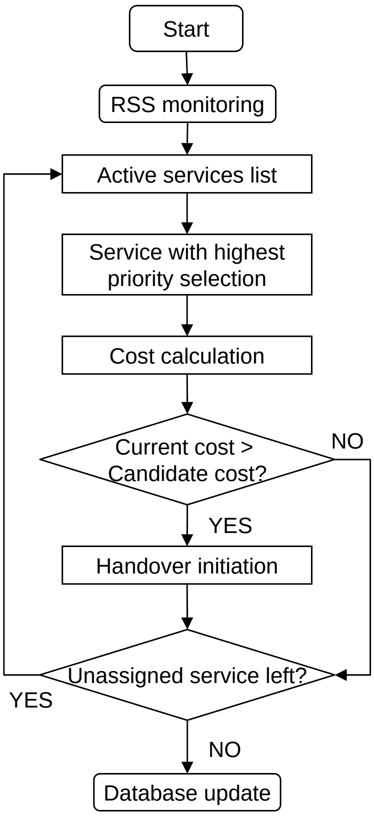

Figure 2.11: Zhu and McNair's VHD heuristic [ZM04, ZM06].

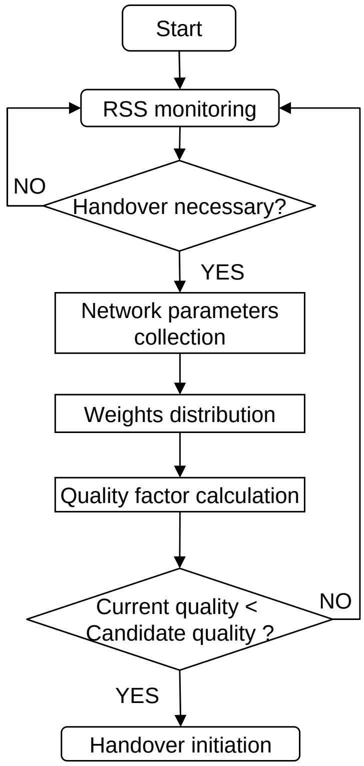

Figure 2.12: Hasswa et al.'s VHD heuristic [HNH06, NHH06b].

service with the highest priority is calculated by

$$C_s^n = \sum W_{s,j}^n Q_{s,j}^n, \quad E_{s,j}^n \neq 0, \tag{2.4.9}$$

where C_s^n is the per-service cost for network n, $Q_{s,j}^n$ is the normalized QoS provided by network n for parameter j and service s, $W_{s,j}^n$ is the weight which indicates the impact of the QoS parameter on the user or the network, and $E_{s,j}^n$ is the network elimination factor, indicating whether the minimum requirement of parameter j for service s can be met by network n. The total cost is the sum of the cost of each QoS parameter, including the bandwidth, battery power and delay. The service is handed over to the network with the minimum cost. The flowchart of this algorithm is shown in Figure 2.11.

The primary benefits brought by the use of a cost function and by independently initiating handovers for different applications are the increased percentage of user satisfied requests and reduced blocking probability. However, the authors did not discuss how the QoS factors are normalized or how the weights for the QoS factors are assigned.

A Cost Function Based Heuristic with Normalization and Weights Distribution

Similar to Zhu and McNair's method [ZM04, ZM06], Hasswa et al. also proposed a cost function based handover decision algorithm in which the normalization and weights distribution methods are provided [HNH06, NHH06b]. A network quality factor is used to evaluate the performance of a handover target candidate as

$$Q_i = \omega_c C_i + \omega_s S_i + \omega_p P_i + \omega_d D_i + \omega_f F_i \tag{2.4.10}$$

where Q_i is the quality factor of network i, C_i, S_i, P_i, D_i and F_i stand for cost of service, security, power consumption, network condition and network performance, and ω_c, ω_s, ω_p, ω_d and ω_f are the weights of these network parameters. Since each network parameter has a different unit, a normalization procedure is used and the normalized quality factor for network n is calculated as

$$\begin{aligned} Q_i \;=\; & \frac{\omega_c(1/C_i)}{\max((1/C_1), ..., (1/C_n))} + \frac{\omega_s S_i}{\max(S_1, ..., S_n)} + \frac{\omega_p(1/P_i)}{\max((1/P_1), ..., (1/P_n))} + \\ & \frac{\omega_d D_i}{\max(D_1, ..., D_n)} + \frac{\omega_f F_i}{\max(F_1, ..., F_n)}. \end{aligned} \tag{2.4.11}$$

A handover necessity estimator is also introduced to avoid unnecessary handovers. Figure 2.12 depicts the operation of this algorithm.

High system throughput and user's satisfaction can be achieved by introducing

Hasswa's heuristic, however, some of the parameters such as security and interference levels are difficult to estimate, and the authors have yet to provide information on how to measure these parameters.

A Weighted Function Based Heuristic

Tawil et al. [TPS08] presented a weighted function based VHD algorithm which delegates the VHD calculation to the visited network instead of the mobile terminal. The weighted function of a network candidate is defined as

$$Q_i = W_B B_i + W_{D_P} \frac{1}{D_{P_i}} + W_C \frac{1}{C_i} \tag{2.4.12}$$

where Q_i represents the quality of network i, B_i, D_{P_i} and C_i are bandwidth, dropping probability and monetary cost of service, and W_B, W_{D_P} and W_C are their weights where

$$W_B + W_{D_P} + W_C = 1. \tag{2.4.13}$$

The network candidate with the highest Q_i is selected as the handover target. The process of this algorithm is shown in Figure 2.13.

By assigning the calculation to the visited network, the resource of the mobile terminal can be saved so that the system is able to achieve short handover decision delay, low handover blocking rate and high throughput. However, the method requires extra cooperation between the mobile terminal and the point of attachment of the visited network, which may cause additional latency and excessive load to the network when there is a large number of mobile terminals.

A summary of the cost function based VHD heuristics is shown in Table 2.3.

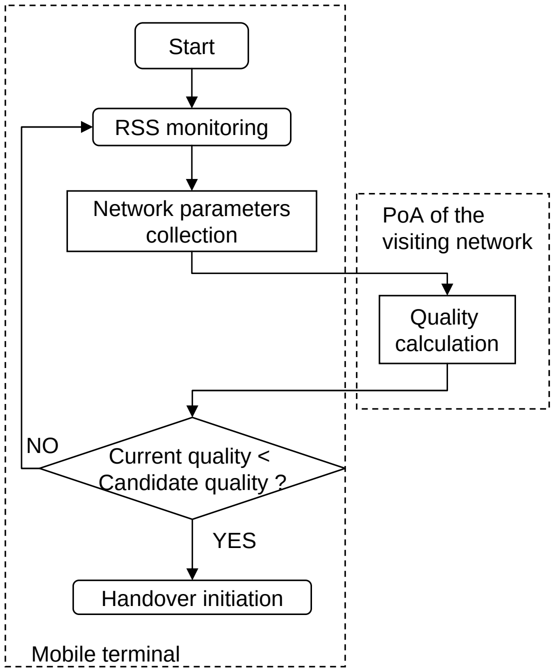

Figure 2.13: Tawil et al.'s VHD heuristic [TPS08].

2.4.4 Combination Algorithms

Combination algorithms are based on artificial neural networks or fuzzy logic, and combine various parameters in the handover decision such as the ones used in the cost function algorithms. Many combination algorithms have been proposed [CSH+01, Zha04, GZX05, HO06, LTD06, NGAM07, PKH+00, XJH07]. In the following sections, three typical combination algorithms are analyzed and evaluated.

A Multilayer Feedforward Artificial Neural Network Based Heuristic

Nasser et al. developed a VHD algorithm based on artificial neural networks (ANN) [NGAM07]. As shown in Figure 2.14, the mobile device collects features of available wireless networks and sends them to a middleware called vertical handover manager through the existing links. These network features are used

Heuristic	Applicable Area	Feature	Advantages	Disadvantages
Zhu and McNair's heuristic [61]	Between any two heterogeneous wireless networks	• A cost function is introduced and users' active applications are individually handed over to target networks with the minimum costs	• Increased user satisfaction • Low blocking probability	• Missing detailed information such as normalization method and weights assignment to make the algorithm realistic
Hasswa et al.'s heuristic [22]	Between any two heterogeneous wireless networks	• Normalization and weights distribution methods are provided • A handover necessity estimator is proposed	• High throughput • High users' satisfaction	• Difficulty in estimating parameters such as security and interference level
Tawil et al.'s heuristic [48]	Between any two heterogeneous wireless networks	• A weighted function is introduced • The handover calculation is delegated to the visiting network instead of the MT	• Short handover delay • Low handover blocking rate • High throughput	• Requirement of cooperation between the MT and the PoA of the visiting network

Table 2.3: A summary of cost function based VHD algorithms.

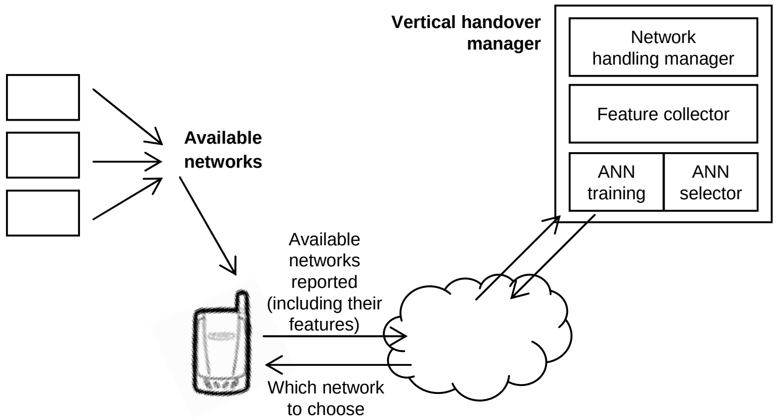

Figure 2.14: Architecture of Nasser et al.'s system [NGAM07].

to help with handover decisions and include network usage cost, network security, network transmission range and network capacity. The vertical handover manager consists of three main components: network handling manager, feature collector and ANN training/selector. A multilayer feedforward ANN is used to determine the best handover target wireless network available to the mobile device, based on the user's preferences.

The topology of the ANN is shown in Figure 2.15. It consists of an input layer,

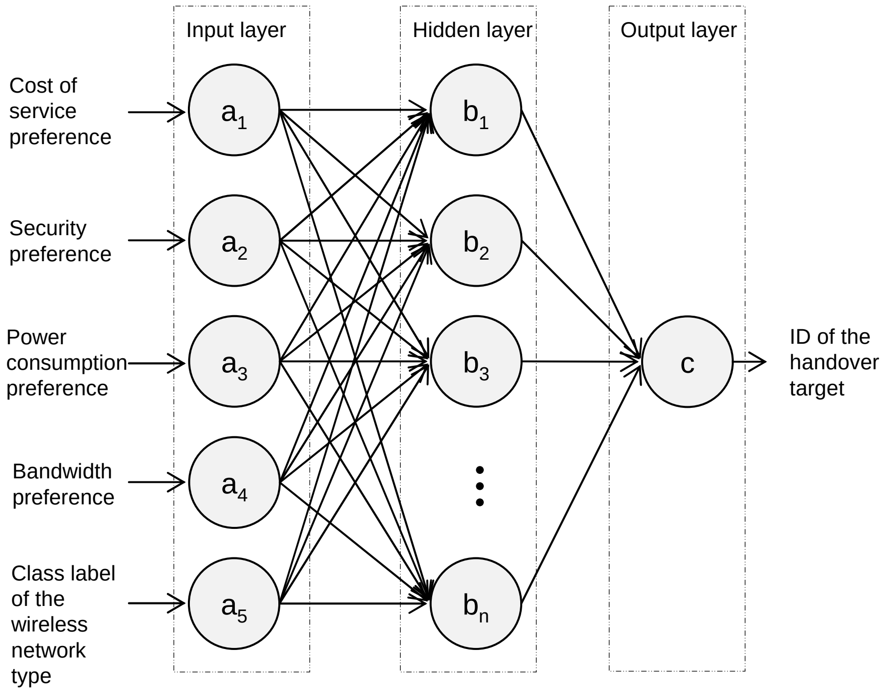

Figure 2.15: Topology of the feedforward ANN used in Nasser et al.'s VHD scheme [NGAM07].

a hidden layer and an output layer. The input layer consists of five nodes representing various parameters of the handover target candidate networks. The hidden layer consists of variable number of nodes which are activation functions. The output layer has one node which generates the network ID of the handover target. All the neurons use sigmoid activation function [SOMGCV+03].

The authors have adopted the same cost function as in [HNH06], and for ANN training they have generated a series of user preference sets with random weights. Then the system has been trained to select the best network among all the candidates.

The authors report that by properly selecting the learning rate and the acceptable error value, the system was able to find the best available network successfully. However, the algorithm suffers from a long delay during the training process.

A Method That Uses Two Neural Networks

Pahlavan et al. [PKH+00] proposed two neural network based decision methods for horizontal and vertical handovers. Here, only the vertical handovers mechanism is discussed.

In the method for vertical handovers, an ANN is used for handovers from the WLAN to the General Packet Radio Service (GPRS). The ANN, as shown in Figure 2.16, consists of an input layer, two middle layers and an output layer. Mobile node performs periodical RSS measurements, and five most recent RSS samples of the access point are fed into the ANN. The output is a binary signal: The value '1' leads to a handover to the GPRS, and the value '0' means that the mobile node should remain connected to the access point.

The ANN is trained before used in the decision process. Training is done by taking a number of RSS samples from the access point and, using a pattern recognition technique, selecting the most suitable network, while minimizing the handover delay and ping-pong effect.

This heuristic can reduce the number of handovers by eliminating the ping-pong effect, but the paper lacks detail on how the neural network is trained and why the particular parameters are selected. This algorithm also has the disadvantage of the increased algorithm complexity and the training process to be performed beforehand.

A Fuzzy Logic Based Heuristic

Besides artificial neural networks, fuzzy logic [HO06, LTD06, XJH07] is also used for creating schemes to deal with a rich set of input parameters for making vertical handover decisions. Xia et al.'s method [XJH07] is a good representative

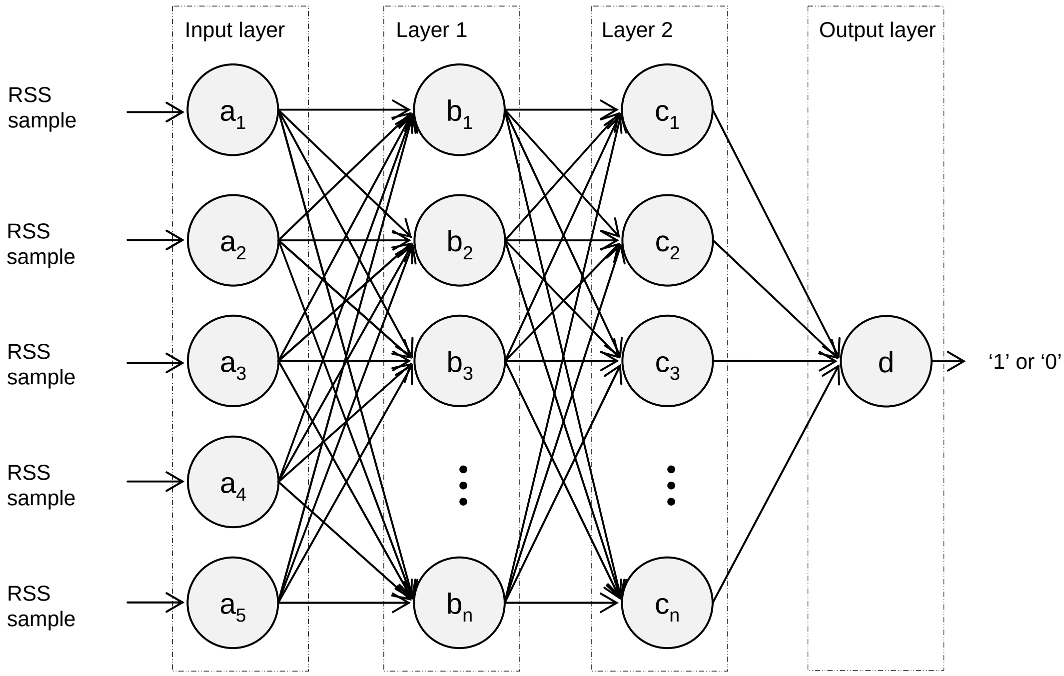

Figure 2.16: Structure of the ANN used for vertical handover decisions in Pahlavan et al.'s VHD scheme [PKH+00]. RSS measurements are done periodically and most recent five samples are fed into the ANN.

example of this approach. This scheme is used to handle handovers between WLANs and UMTS. A pre-decision unit is used in this scheme. In this algorithm, if the mobile terminal is connected to the WLAN, if the velocity of the mobile terminal (v) is higher than a velocity threshold (v_T), a handover to the UMTS is directly initiated to prevent a connection breakdown. Otherwise, the pre-decision unit checks whether the predicted RSS satisfies its requirements. If the predicted RSS from the WLAN (PR_W) is larger than its threshold (P_{rW}), or the predicted RSS from the UMTS (PR_U) is smaller than its threshold (P_{rU}), no handover is triggered. After the pre-decision, the fuzzy logic based normalized quantitative decision (FNQD) is applied. The FNQD has three procedures: fuzzification, normalization and quantitative decision. The three inputs, current RSS, predicted

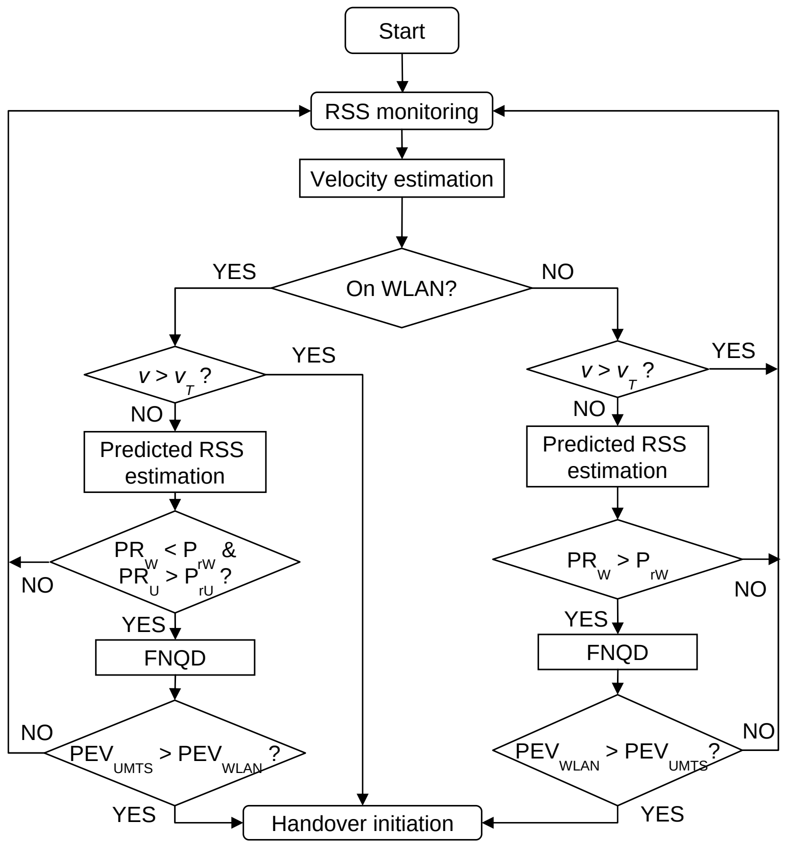

Figure 2.17: Xia et al.'s VHD heuristic [XJH07].

RSS and bandwidth, are fuzzified and normalized to generate performance evaluation values (PEV), and the VHD is made by comparing PEVs of the network candidates.

If the mobile terminal is connected to the UMTS and the WLAN connectivity is available, the pre-decision unit is used to eliminate unnecessary handovers when the velocity of the mobile terminal is larger than the threshold (v_T). A similar process is executed as the one described in the handover from WLANs to UMTSs. The process of this algorithm is illustrated in Figure 2.17.

The heuristic in this study is able to achieve improved performance by reducing the number of unnecessary handovers and avoiding the ping-pong effect. However, when the PEVs are calculated, fixed weights are assigned to the three inputs.

Heuristic	Applicable Area	Feature	Advantages	Disadvantages
Nasser et al.'s heuristic [36]	Between any two heterogeneous wireless networks	A cost function is adopted and the system is trained before being used in the handover decision	• High success rate in finding the best network candidate	• Training delay • Increased system complexity
Pahlavan et al.'s heuristic [40]	From WLANs to GPRS	The RSS samples are collected as the inputs of the neural network and the system is trained before being used in the handover decision	• Reduced number of handovers • Elimination of the ping-pong effect	• Lack of detail on training process and parameters selection • Training delay • Increased system complexity
Xia et al.'s heuristic [52]	Between WLANs and cellular networks	Current RSS, predicted RSS and bandwidth are fuzzificated and normalized to be used as the handover decision criteria	• Reduced number of handovers • Elimination of the ping-pong effect	• Fixed weights which fail to meet the need of continuously changing wireless environment

Table 2.4: A summary of combination algorithms.

This is not practical because the network condition and user requirements vary in different situations. In addition, more performance evaluation criteria such as handover delay and system load need to be addressed.

A summary of the combination VHD heuristics is shown in Table 2.4.

2.5 Comparison of the Approaches

2.5.1 Comparison Among Existing VHD Algorithms

So far twelve VHD algorithms have been discussed and classified into four groups based on the criteria they use for making handover decision. To provide an overall comparison of the four groups, their features are summarized on five aspects: networking technologies that they can be applicable, input parameters, handover target selection criteria, complexity and reliability in Table 2.5.

The applicable network technologies for RSS based VHD algorithms are usually between macrocellular and microcellular networks, e.g. 3G and WLANs. The algorithms tend to make full usage of WLANs because of the availability of high

Group	Applicable networking technologies	Input parameters	Handover target selection criteria	Complexity	Reliability
RSS based VHD algorithms	Usually between macrocellular and microcellular networks	RSS as the main input	The network candidate with the most stable RSS	Simple	Reduced reliability because of the fluctuation of RSS
Bandwidth based VHD algorithms	Between any two heterogeneous networks	Bandwidth combined with other parameters such as RSS	The network candidate with the highest bandwidth	Simple	Reduced reliability because of the changing available bandwidth
Cost function based VHD algorithms	Between any two heterogeneous networks	Various parameters such as cost, bandwidth and security	The network candidate with the highest overall performance	Complex	Reduced reliability because of the difficulty in measuring some parameters
Combination algorithms	Between any two heterogeneous networks	Different input parameters depending on different methods	The network candidate with the highest overall performance	Very complex	High reliability because of the training of the system

Table 2.5: A comparative summary of the four groups.

bandwidth provided and their cost efficiency. The other three types of VHD algorithms can be applied for handovers over all kinds of wireless networks.

As for the input parameters, RSS is used as the main input in RSS based VHD algorithms, while the RSS combined with the bandwidth information is usually adopted in bandwidth based VHD algorithms. Various network parameters are used in cost function based or combination algorithms, such as monetary cost, bandwidth, security and power consumption.

For handover target selection criteria, the candidate network with the most stable RSS and highest bandwidth is selected as the handover target in RSS and bandwidth based VHD algorithms, respectively. On the other hand, combination or cost function based algorithms attempt to choose the target network with the highest overall performance. The overall performance is calculated based on the various network parameters.

From the complexity point of view, among the four groups, RSS based algorithms are usually the simplest, followed by the bandwidth based algorithms. Cost function based VHD algorithms tend to be more complex as they need to collect and normalize various network parameters, and combination algorithms are the most challenging ones because of their pre-training requirements.

Finally, reliability varies among the algorithms. Fluctuations of RSS decreases the reliability of RSS based VHD algorithms, and the difficulty of available bandwidth measurements reduces the reliability of bandwidth based VHD algorithms. In cost function based algorithms, some parameters such as security level are hard to measure, and thus the reliability of these algorithms degrade. As for combination algorithms, since the systems are trained beforehand, they can be considered as the most reliable among the four groups.

For a better understanding of the performance of different VHD algorithms, a quantitative comparison is provided based on the performance metrics mentioned in Section 2.3.3. Since the authors of each algorithm provide different performance parameters in their studies, direct comparisons are impossible. In Table 2.6, a summary quantitative comparison is provided based on four performance parameters: delay, number of handovers, handover failure probability and throughput, based on the information provided in the papers. As can be seen, relatively high delays occur by using RSS based algorithms proposed in [YMŞ08] and [ZLS06], while the authors in [LCCS05] and [TPS08] argue that their bandwidth and cost function based algorithms are able to maintain shorter handover delays. Combination algorithms suffer from the longest delay among the four groups because of the system complexity. For the case of number of handovers, the use of algorithms in [YMŞ08] and [ZLS06] lead to reduced number of handovers, the algorithm in [YGQD07] introduces excessive handovers because of the variation of SINR, the algorithm in [CCHL07] is able to keep the unnecessary handover probability at a low level, and algorithms in [PKH+00] and [XJH07] reduce the number of handovers by eliminating the ping-pong effect. Handover failure probability can always be kept under the desirable value for algorithms in [MA06] and [YMŞ08], while high handover failure probability is observed for the algorithm in [CCHL07] without inclusion of RSS. The algorithm in [TPS08] can achieve low failure rate due to its distribution of the decision calculation. As for the throughput, bandwidth and cost function based algorithms are able to achieve higher throughput than RSS based algorithms, unfortunately the throughput of combination algorithms are not provided by the authors.

In summary, RSS and bandwidth based VHD algorithms are usually simple, but they only consider one or two handover criteria as the inputs and other important parameters such as monetary cost or power consumption level of the networks are ignored. Furthermore, they are usually targeted to only two specific types of network technologies. Cost function based and combination algorithms are more complex, and they take into account a wider range of network parameters as compared to others. However, they are mostly in the theoretical analysis stage or are too complicated for implementation yet.

2.6 Conclusions

In this chapter, a comprehensive survey of VHD algorithms is presented. These algorithms are categorized into four groups: RSS, bandwidth, cost function and combination based. VHD algorithms in the published research literature lack a comprehensive consideration of various network parameters, user mobility and user preferences. The research project presented in this thesis focuses on this issue, and provides an integrated solution to the optimization of the VHD process. In the next chapter, the framework of the proposed VHD scheme is provided.

Groups/Heuristics		Delay	Number of Handovers	Handover Failure Probability	Throughput
RSS based	Zahran et al.'s algorithm [59]	Relatively high packet delay probability (up to 1%) but can be reduced by adjusting ASST	Reduces up to 85% comparing with traditional hysteresis VHD	Not provided	Decreases as the velocity increases; Can provide overall higher throughput (up to 33%) than traditional hysteresis VHD)
	Mohanty and Akyildiz's algorithm [33]	Not provided	Not provided	Can be always kept under the desirable value (2%) as the velocity increases	Not provided
	Yan et al.'s algorithm [53]	Extra RSS sampling delay (up to 2s)	Decreases as the velocity increases; The unnecessary handover probability can be always kept under the desirable value (0.04)	Can be always kept under the desirable value (0.02) as the velocity increases	Not provided
Bandwidth based	Lee et al.'s algorithm [26]	Short handover delay (average 455ms) achieved by considering application types	Not provided	Not provided	Higher throughput (up to 400%) than the traditional method in the handover period
	Yang et al.'s algorithm [56]	Not provided	Excessive handovers can be introduced because the variation of SINR	Not provided	Higher overall throughput (up to 40%) than RSS-based handover algorithms
	Chi et al.'s algorithm [14]	Not provided	Small unnecessary handover probability (up to 1.5%)	High handover failure probability without considering RSS	High throughput achieved by balancing the traffic load
Cost function based	Zhu and McNair's algorithm [61]	Not provided	Not provided	Not provided	High overall throughput achieved by spreading users' services over several networks
	Hasswa et al.'s algorithm [22]	Not provided	Not provided	Not provided	Increases by up to 57.9% in different background traffic
	Tawil et al.'s algorithm [48]	Around 50% shorter handover delay compared to centralized VHD	Not provided	Low handover failure rate due to the distribution of the decision calculation	Around 17% higher throughput compared to centralized VHD
Combination algorithms	Nasser et al.'s algorithm [36]	Long handover delay because of the training needed	Not provided	Not provided	Not provided
	Pahlavan et al.'s algorithm [40]	Long delay because of the increased complexity and the training	Reduced number of handovers by eliminating the ping-pong effect	Not provided	Not provided
	Xia et al.'s algorithm [52]	Not provided	Reduced number of handovers by eliminating the ping-pong effect	Not provided	Not provided

Table 2.6: A comparative summary of the twelve VHD algorithms presented in this survey.

Chapter 3

Overview of VHD Optimization Scheme

3.1 Introduction

This chapter introduces the overall framework of the proposed VHD Optimization scheme. The framework consists of three main components: handover necessity estimation (HNE) module, handover target selection (HTS) module and handover triggering condition estimation (HTCE) module.

3.1.1 A Use Case Scenario

To provide a better understanding of the framework of the handover optimization method, an example is given to demonstrate a scenario of the chronological order in which a MT invokes the three modules. Figure 3.1 shows this scenario which is described as follows:

1. The MT is connected to a 3G network and senses the availability of a WLAN (WLAN1), so it invokes the HNE module to decide whether to handover to the WLAN or not.

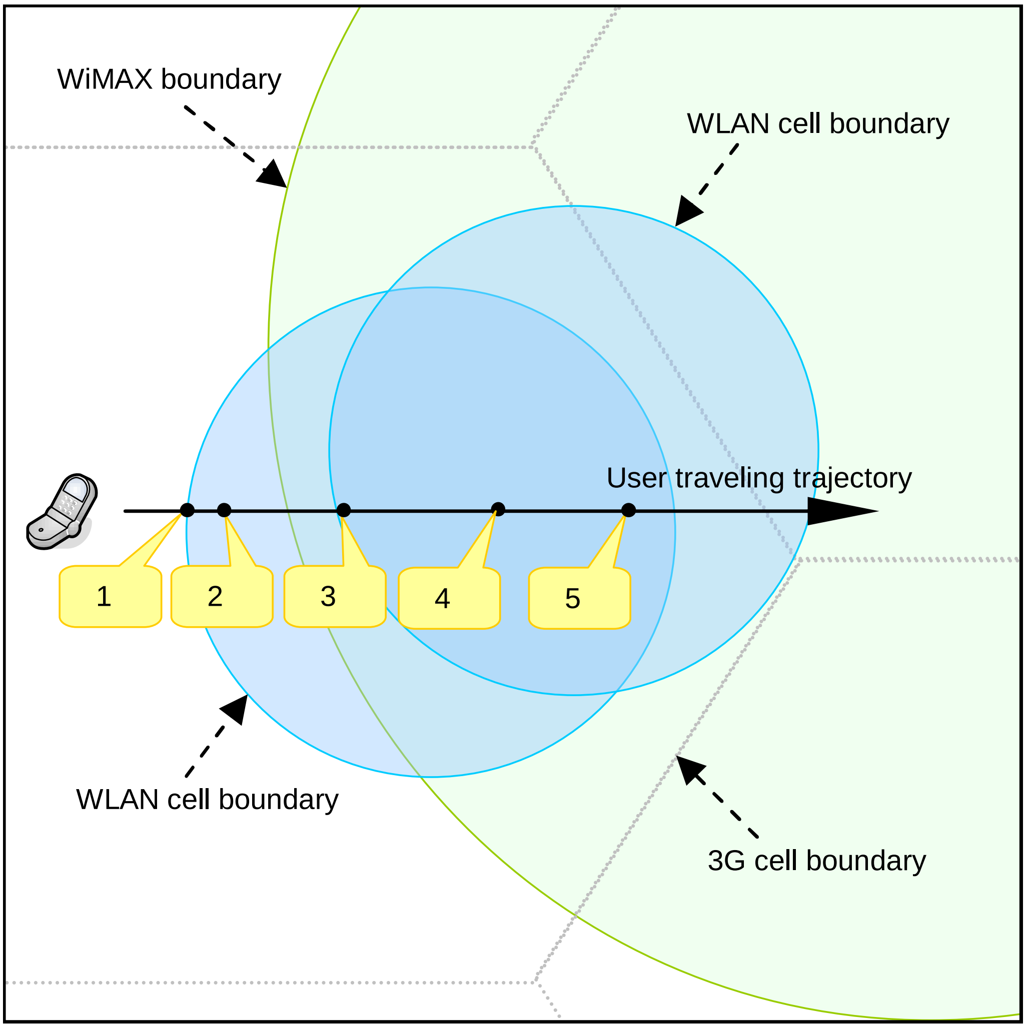

Figure 3.1: Vertical handover decision: a user scenario.

2. The handover is necessary, so the MT performs a handover to the WLAN.

3. The MT senses the availability of multiple candidate networks including a second WLAN (WLAN2) and a WiMAX network. It then invokes the HNE unit again to evaluate the necessity of a handover to these candidate networks, and find both are necessary. Afterwards, the HTS module is invoked to select the handover target network, and WLAN2 is chosen.

4. The MT senses that the RSS from the WLAN continuously deteriorates, which invokes the HTCE module to find the triggering condition to handover out of WLAN2.

5. Once the handover triggering condition estimated in step 4 is satisfied, a handover to the 3G network is triggered.

3.1.2 The Handover Process

The process of the proposed VHD method is shown in Figure 3.2. As shown in the figure, the handover decision method is initiated either because a more preferable PoA becomes available or the RSS from the current PoA falls below a given threshold. The handover decision process under these two conditions is explained below.

If one or more preferable PoAs become available:

Step 1. The handover necessity estimation unit determines the necessity of triggering a handover for each candidate network.

Step 2. If a handover is not necessary, the handover decision process is terminated; if a handover to one or more candidate networks is preferred, handover decision parameters of these candidate networks are collected. A cost function that assigns different weights to these parameters is then used to assign a cost of handover to each of the candidate networks.

Step 3. A handover to the candidate network with the minimum cost is immediately initiated.

If the user is moving out of the current PoA and thus the RSS of the current PoA continuously deteriorates:

Step 1. The handover time estimation unit calculates the time for triggering a handover.

Step 2. If only one candidate network is available, a handover to this network

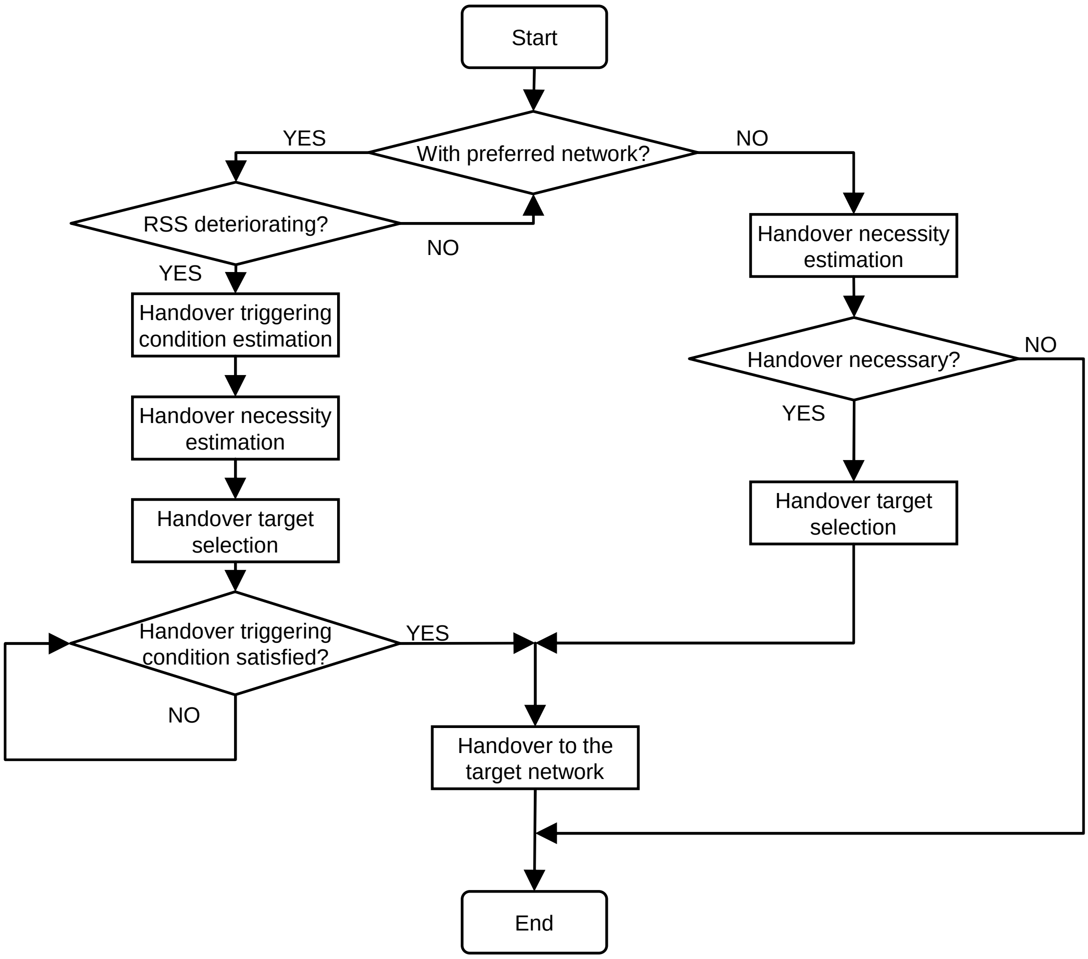

Figure 3.2: Vertical handover decision process.

is initiated at the time calculated in Step 1 and the handover decision process is then terminated; if more than one network are available, the algorithm for one or more available PoAs (described above) is invoked.

Step 3. A handover to the candidate network with the minimum cost is initiated at the time calculated in Step 1.

In the following sections, we discuss the three units of the VHD framework in detail.

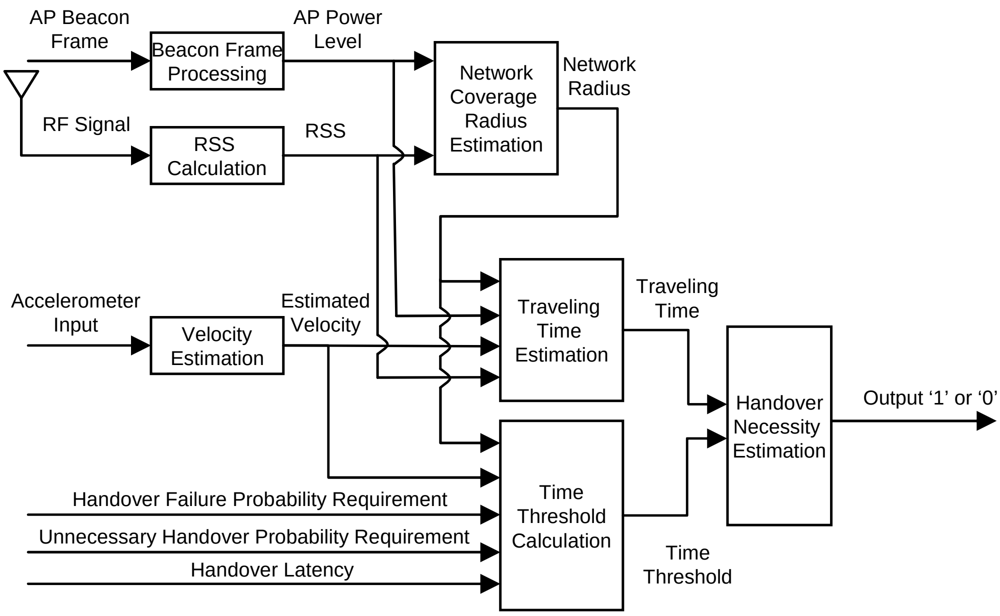

Figure 3.3: Block diagram of handover necessity estimation (HNE) unit.

3.2 Handover Necessity Estimation

The handover necessity estimation (HNE) unit determines the necessity of making a handover to an available network. HNE takes various network parameters as its inputs and generates a binary value as its output. The inputs include: the AP power level, RSS samples, the radius of the network, the velocity of the MT, the handover latency, and the handover failure and unnecessary handover probability requirements. An output of '1' means a handover is necessary, and an output of '0' means the handover is not necessary. The block diagram of HNE is shown in Figure 3.3.

HNE consists of two units, traveling time estimation and time threshold calculation. The condition of triggering the HNE process is that a more preferable network is available, e.g. a WLAN coverage becomes available while the MT is currently connected to a cellular network. Once the process is triggered, the HNE

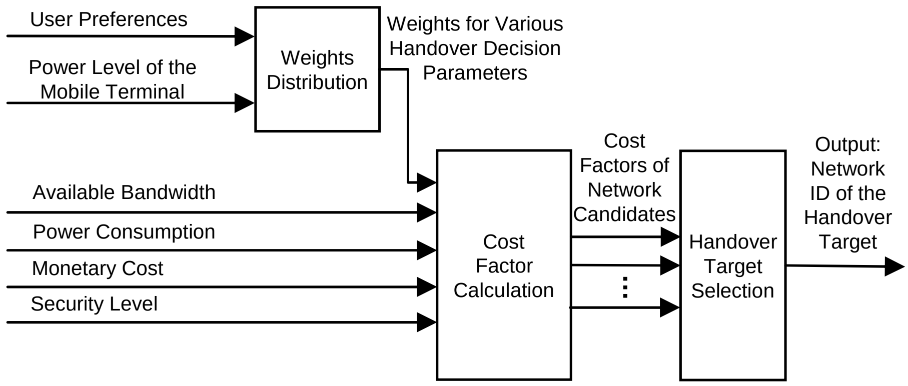

Figure 3.4: Block diagram of handover target selection (HTS) unit.

unit estimates the traveling time inside a candidate network and calculates a time threshold. The traveling time is estimated by using RSS samples collected by the MT, the power level and radius of the AP of the network, and the velocity of the MT. The time threshold is calculated based on the radius of the AP, the velocity of the MT, the handover latency, and the handover failure and unnecessary handover probability requirements. HNE then compares the estimated traveling time against the time threshold: if the traveling time is greater than the threshold, an output of '1' is generated and a handover is necessary; otherwise an output of '0' is generated and a handover is unnecessary. HNE carries out this process for all the candidate networks and generates a binary output for each of them.

More details of the algorithms used in HNE are provided in Chapter 4.

3.3 Handover Target Selection

The handover target selection (HTS) unit selects the "best" candidate network among the available candidate networks. HTS adopts a cost function to calculate the cost factor of each candidate network, and chooses the network with the

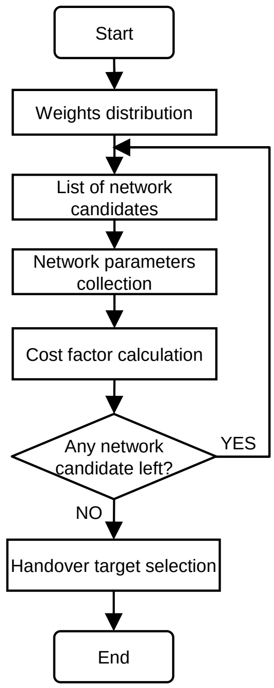

Figure 3.5: Handover target selection process.

minimum cost as the handover target. The cost function considers various network parameters such as the available bandwidth, the power consumption, the monetary cost and the security level. A weight distribution unit is also included to assign different weights to the parameters based on user preferences and the power level of the MT. A block diagram of HTS is shown in Figure 3.4.

HTS consists of two units, weights distribution and cost factor calculation. The process of handover target selection is shown in Figure 3.5 and the steps in this process are further explained below.

Step 1. The weights distribution unit collects user preferences and the power level of the MT and generates weight factors for various handover decision parameters.

Step 2. HTS lists all the available candidate networks in random order.

Step 3. Starting with the candidate network with the highest priority, network parameters for handover decision are collected. These parameters include the available bandwidth, the power consumption, the monetary cost and the security level.

Step 4. A cost factor is calculated using the parameters collected in Step 3 and a cost function.

Step 5. HTS checks whether there is a candidate network left. If there is, the process returns to Step 3; otherwise the process moves to Step 6.

Step 6. The network with the minimum cost factor is selected as the handover target.

The algorithms used in the weights distribution and cost factor calculation units are explained in Chapter 5.

3.4 Handover Triggering Condition Estimation

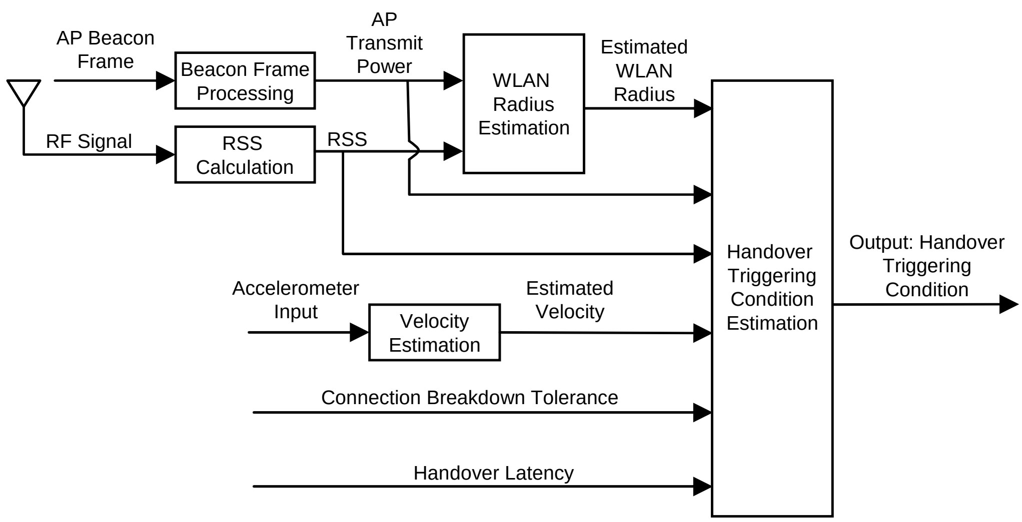

Figure 3.6: Block diagram of handover triggering condition estimation (HTCE) unit.

The handover triggering condition estimation (HTCE) unit determines a proper time to initiate a handover out of the current connected network to prevent connection breakdowns as well as to maximize the usage of the preferable network. HTCE takes inputs including AP power level, RSS samples, the radius of the current connected network, the velocity of the MT, the handover latency and the connection breakdown probability requirement, and generates the handover triggering condition as its output. A block diagram of HTCE is shown in Figure 3.6.

The condition of triggering the HTCE process is that the RSS from the current PoA continuously deteriorates and another PoA is available. Once the process is triggered, the HTCE unit estimates a RSS threshold, which is calculated based on the RSS samples collected by the MT, the power level and radius of the AP, the velocity of the MT, the handover latency, and the connection breakdown probability requirement. A handover is triggered when the RSS from the current PoA drops below the RSS threshold.

More details of the algorithms used in HTCE are provided in Chapter 6.

3.5 Conclusions

In this chapter, the framework of the proposed handover decision method is presented. The handover process, along with the three main components of the framework, handover necessity estimation, handover triggering condition estimation and handover target selection, have been described. The objective of this framework is to provide an optimized handover decision which

1. minimizes unnecessary handovers and handover failures using the handover necessity estimation unit,
2. maximizes the users' satisfaction using the handover target selection unit, and

3. keeps the connection breakdown probability below desirable limits, and provides the user with control over the tradeoff between connection breakdown probability and WLAN usage.

In the next three chapters, the algorithms used in handover necessity estimation, handover triggering condition estimation and handover target selection are explained in detail.

Chapter 4

A New Method for Handover Necessity Estimation

4.1 Introduction

In order to maintain seamless user roaming and optimize network resource usage, it is desirable to minimize handover failures and unnecessary handovers. As discussed in Section 2.4.1, Mohanty [Moh06] presented an algorithm for calculating a boundary area based on the speed of the MT and the WLAN cell size. In this algorithm, a handover from a WLAN to a 3G network is triggered when the MT enters the boundary area of the WLAN and handover procedures are completed before the MT leaves the WLAN. This algorithm operates efficiently for handovers from WLAN to 3G as it reduces the handover failure probability.

However, in the mobility architecture using this algorithm, and also in most of the other handover decision methods such as in [VRWF03], handovers from the cellular network to the WLAN are initiated once the MT enters the WLAN coverage area. This is not effective enough in situations where the MT travels through an area close to the coverage boundary of the WLAN at speeds above a certain threshold, since handovers to the WLAN become unnecessary. It is always better to avoid these handovers as much as possible since they lead to network resource wastage [CS05]. Furthermore, if the handover process has not been completed before the MT leaves the WLAN coverage area, connection breakdown inevitably occurs. In the method presented in [Moh06] for handovers from the cellular network to the WLAN, the MT remains connected to both networks while staying in a boundary cell of the WLAN in order to avoid connection breakdown and also the ping-pong effect. However, this approach does not take into consideration the network resource wastage caused by unnecessary handovers. As yet, few studies on handover necessity estimation or on efficient methods for minimizing unnecessary handovers has been presented.

In this chapter, a handover necessity estimation method is introduced. This method is devised for minimizing handover failures and unnecessary handovers from cellular networks to WLANs, by estimating the necessity of a handover. The estimation involves two steps: traveling time prediction and time threshold calculation.

This chapter is organized as follows: Section 4.2 presents the traveling time prediction algorithm, and Section 4.3 introduces the time threshold calculation algorithm.

4.2 Traveling Time Prediction

In this section, the traveling time prediction algorithm is presented. Section 4.2.1 provides the mathematical justification of the algorithm, and Section 4.2.2 discusses the accuracy of the time prediction and provides a way of improving this accuracy.

4.2.1 Traveling Time Prediction Using RSS Measurements and speed Information

The handover necessity estimation relies on an algorithm which attempts to predict the traveling time in a WLAN cell coverage area by using successive RSS measurements. The algorithm works under the following assumptions:

- the WLAN cell has a circular geometry;
- the MT travels through the WLAN cell coverage in a straight line with a constant speed; and
- the propagation environment in the WLAN coverage is modeled using the log-distance path loss model [Stü01].

Figure 4.1 shows the traveling time prediction scenario. The relationship between RSS (in dBm), and the distance between the AP and the MT at any point P inside the WLAN coverage area is obtained by using the log-distance path loss model:

$$\mathrm{RSS}_P = P_{\mathrm{Tx}} - \mathrm{PL}_{\mathrm{ref}} - 10\beta \log_{10} \frac{l_{\mathrm{OP}}}{d_{\mathrm{ref}}} + X_\sigma, \tag{4.2.1}$$

where P_{Tx} is the transmit power of the WLAN AP in dBm, l_{OP} is the distance between the AP and point P, d_{ref} is the distance between the AP and a reference point, $\mathrm{PL}_{\mathrm{ref}}$ is the path loss at the reference point in dB, β is the path loss exponent, and X_σ is a Gaussian distributed random variable with a mean of zero and a standard deviation σ in dB.

Estimation of the traveling time of the MT by using RSS measurements is done in the following way. It is assumed that the MT starts receiving sufficiently strong signals (i.e., it “enters” the WLAN cell) at point P_i and the signal strength drops below the usable level at point P_o, and M is the middle point of the traveling

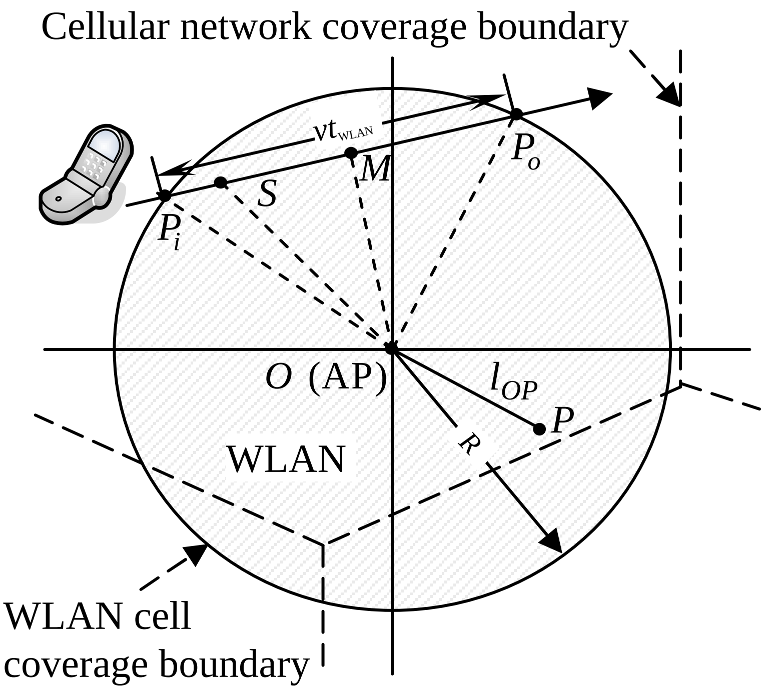

Figure 4.1: A scenario for traveling time prediction in a WLAN cell.

trajectory, as shown in Figure 4.1. By using (4.2.1), when the MT enters the WLAN cell coverage area (i.e., the RSS level detected by the MT reaches a pre-determined threshold) at time t_{P_i}, the distance l_{OP_i} (an approximate value for the cell radius R) can be calculated using

$$R \cong l_{OP_i} = d_{\text{ref}} 10^{\frac{P_{\text{Tx}} - \text{PL}_{\text{ref}} - \text{RSS}_{P_i}}{10\beta}}, \tag{4.2.2}$$

where RSS_{P_i} is the RSS at the entry point P_i.

To estimate the traveling time t_{WLAN}, the MT takes another RSS sample at point S at time t_S. Using Equation (4.2.1), the distance between O and S, l_{OS}, is estimated by

$$l_{OS} = d_{\text{ref}} 10^{\frac{P_{\text{Tx}} - \text{PL}_{\text{ref}} - \text{RSS}_S}{10\beta}}, \tag{4.2.3}$$

where RSS_S is the RSS at the sampling point S.

From the geometric configuration of Figure 4.1, the following equations are obtained:

$$l_{P_iM}^2 + l_{OM}^2 = l_{OP_i}^2 = R^2 \tag{4.2.4a}$$

$$l_{OM}^2 + l_{SM}^2 = l_{OS}^2 \tag{4.2.4b}$$

$$l_{SM} = l_{P_iM} - l_{P_iS}, \tag{4.2.4c}$$

where l_{P_iM}, l_{OM}, l_{SM}, l_{OS} and l_{P_iS} are the distances between the entry point P_i and the middle point M, the AP location O and point M, the sampling point S and point M, points O and S, and points P_i and S, respectively.

By substituting Equation (4.2.4c) in Equation (4.2.4b), the following equation is obtained:

$$(l_{P_iM} - l_{P_iS})^2 + l_{OM}^2 = l_{OS}^2. \tag{4.2.5}$$

Let v be the speed of the MT, which is a constant during the time period when the MT crosses the WLAN cell coverage t. Thus

$$l_{P_iM} = \frac{vt}{2} \tag{4.2.6a}$$

$$l_{P_iS} = v(t_S - t_{P_i}), \tag{4.2.6b}$$

where t_S and t_{P_i} are times at sampling and entry points S and P_i, respectively.

By substituting Equation (4.2.6) in Equations (4.2.4a) and (4.2.5), the following equations are obtained:

$$\left(\frac{vt}{2}\right)^2 + l_{OM}^2 = R^2 \tag{4.2.7a}$$

$$\left[\frac{vt}{2} - v(t_S - t_{P_i})\right]^2 + l_{OM}^2 = l_{OS}^2. \tag{4.2.7b}$$

Based on Equation (4.2.7), an estimate of traveling time t_{WLAN} is calculated as

$$t_{\text{WLAN}} = \frac{R^2 - l_{OS}^2 + v^2(t_S - t_{P_i})^2}{v^2(t_S - t_{P_i})}. \tag{4.2.8}$$

Substituting Equations (4.2.2) and (4.2.3) in Equation (4.2.8), the ultimate equation of t_{WLAN} is

$$t_{\text{WLAN}} = \frac{d_{\text{ref}} 10^{\frac{2\left(P_{\text{Tx}} - \text{PL}_{\text{ref}} - RSS_{P_i}\right)}{10\beta}} - d_{\text{ref}} 10^{\frac{2\left(P_{\text{Tx}} - \text{PL}_{\text{ref}} - RSS_S\right)}{10\beta}} + v^2(t_S - t_{P_i})^2}{v^2(t_S - t_{P_i})}. \tag{4.2.9}$$

The traveling speed of the MT v is measured by an accelerometer embedded in the MT [ZL07]. Accelerometers can be used in handsets for various purposes and one purpose is to accurately estimate the speed of the MT.

4.2.2 The Impact of Fading Phenomena on RSS Measurements and Possible Solutions

The presented traveling time prediction algorithm mainly relies on RSS measurements. The RSS fluctuates because of the fading phenomena, and since such fluctuations inevitably affect the accuracy of the time prediction method and consequently affect the performance of the VHD method, it is necessary to investigate the causes of these fluctuations and solutions to compensate for them.

Fading in mobile radio systems is divided into two different types, fast and slow fading [Stü01]. Fast fading (fast fluctuations in the received signal's amplitude, phase and angle of arrival) happens when a signal travels from transmitter to receiver over multiple paths caused by propagation mechanisms. This type of fading is often modeled as a Rayleigh or Rician random variable [GS02]. On the other hand, slow fading, or shadow fading, represents the average signal power fluctuations or path loss due to motion over large areas. This phenomenon is affected by prominent terrain contours (e.g., hills, forests, buildings etc.) between the transmitter and receiver and is usually modeled with a log-normal random variable. For mobile radio applications, the channel is time-variant because motion between the transmitter and receiver results in propagation path changes. Therefore, the mobile radio roaming over a large area must process signals that, in addition to path loss associated with distance, experience two types of fading: small-scale Rayleigh or Rician fading superimposed on large-scale log-normal fading [MP05].

In RSS measurements, fast fading is averaged out and can be neglected due to its short correlation distance [Sin07]. As a result, in this work the only considered phenomenon that affects the accuracy of the time prediction is shadow fading. In order to maintain low complexity, the effect of shadow fading on distance measurement variables is usually assumed as Gaussian statistics [Sin07], and thus in this work the shadow fading is represented by the Gaussian distributed random variable ξ in Equation (4.2.1).

The simulated RSS variation from the MT caused by the shadow fading and change of the distance from the AP is depicted in Figure 4.2. The graph was obtained with the following parameters using Equation (4.2.1): the transmit power of the AP was 27 dBm, the path loss exponent β was 3.2, and the standard deviation of shadowing σ was 4.3 dB, for typical urban environments [H$^+$02].

To reduce the impact of shadow fading, different methods have been proposed in the literature [GG05, Sin07]. We consider the following three solutions to reduce the impact of shadow fading: the spatial RSS digital map, the global positioning system (GPS), and the moving average of RSS samples. Their details are described below.

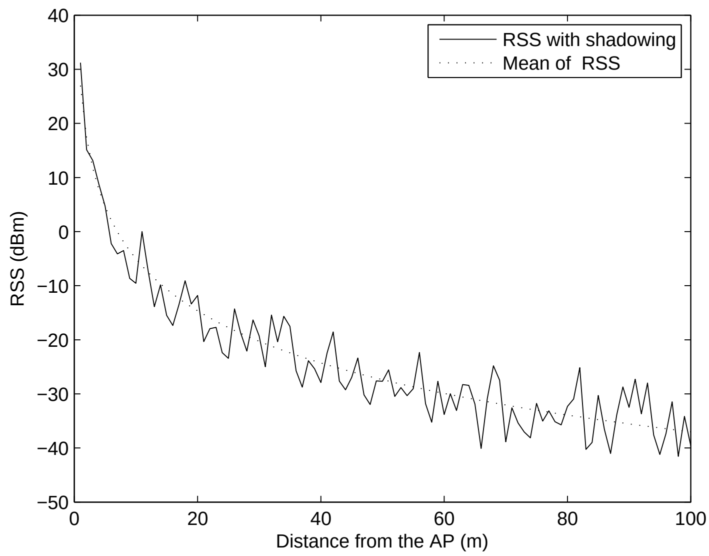

Figure 4.2: Simulated RSS variation from the MT caused by the shadow fading and change of the distance from the AP.

The first solution to the shadow fading is to utilize a predicted or measured spatial digital map with RSS values [GG05]. A digital map contains, for instance, RSS measurements relative to the reference points either predicted or provided via dedicated measurement scans in the service area. The former is conducted in the network deployment phase using graphical information systems dedicated to network planning, and the latter is only plausible in very limited service areas, like in indoor environments. An example of the digital map is provided in [GG05]. This method requires pre-collected RSS information.

The second solution is to utilize the GPS to provide mobility information. The MT is able to obtain the position and coverage information from the local APs. By using this information and the mobility information provided by GPS, the MT would be able to prediction the traveling time within the AP coverage. An example of such is provided in [LSG01]. This method requires MTs with GPS equipments.

The third solution is to collect more RSS samples and average them in a sliding window [Sin07]. In this research, this solution is adopted to reduce the shadow fading effect, because it requires neither pre-collected RSS data nor a GPS equipment which consumes extra battery power, and is able to provide a satisfactory level of accuracy for the traveling time prediction. Its details are described below.

First of all, a sliding window with the size of N data elements can be defined as follows. Data elements arrive at each instant and expire after exactly N time steps; and, the portion of data that is relevant to gathering statistics or answering queries is the set of the last N elements to arrive. The sliding window refers to the window of active data elements at any time instant [DGIM02]. Then using a sliding window with the size of N RSS samples, the measured RSS samples are averaged before they could be used to trigger the handover initiation algorithm as follows [CGS94]:

$$\text{RSS}(k) = \frac{1}{N_w} \sum_{n=0}^{N-1} \text{RSS}(k-n) W_n. \qquad (4.2.10)$$

where RSS(k) is the RSS of the k^{th} sample after averaging, W_n is the weight assigned to the sample taken at the end of the $(k-n)^{\text{th}}$ interval, and $N_w = \sum_{n=0}^{N-1} W_n$. For a rectangular window, equal weight is assigned to all the previous samples in the averaging window, therefore, $W_n = 1$ for all n.

In the simulation experiment, it is observed that shadow fading had greater impact on the accuracy of distance estimations for lower speeds than that for higher speeds of the MT. Thus, at lower speeds larger number of RSS samples are desirable to compensate for higher shadowing effects. The value of N is therefore dynamically adjusted to the MT's speed (v) as

$$N = \frac{N_s}{v}, \qquad (4.2.11)$$

where N_s is an empirical value representing the size of the sliding window when $v = 1m/s$.

The standard deviation of the averaged shadowing samples is calculated as

$$\sigma_a = \sqrt{\frac{\sigma^2}{N}[1 + 2\sum_{n=1}^{N-1}(1 - \frac{n}{N})\rho^n]} \tag{4.2.12}$$

where σ and σ_a are the standard deviation of the shadow fading before and after average of RSS measurements over N samples, and ρ denotes the autocorrelation coefficient of the shadow fading.

By using the averaged RSS measurements, shadow fading effects could be reduced and thus the accuracy of the time prediction could be increased. For example, by using a rectangular sliding window with size 20, the standard deviation of the shadow fading can be reduced from 4.3 dB to 1.17 dB, according to Equation (4.2.12).

4.3 Time Threshold Calculation

In this section, the algorithm used for time threshold calculation is introduced. This algorithm contains two parts which aim to minimize handover failures and unnecessary handovers, respectively, which are both built on probability calculation. They are discussed in separate sections below.

4.3.1 Time Threshold Calculation for Minimizing Handover Failures

The purpose of the time threshold calculation presented in this section is to keep the number of handover failures under a desirable threshold. That is, for example, if the system designer has a requirement of limiting the probability of handover failures under 1%, then the time threshold is adjusted to make the ratio of the number of failed handovers to the total number of handovers below 1%. The time threshold is calculated using mathematical modeling and probability calculation as explained below.

It is assumed that the entry and exit points P_i and P_o can be any arbitrarily chosen points on the circle enclosing the WLAN coverage area, with equal probability (Figure 4.1). Then the angles θ_i and θ_o are both uniformly distributed in $[0, 2\pi]$, and $\theta = |\theta_i - \theta_o|$.

The first step is to calculate the probability density function (PDF) of θ.

The PDFs of the locations of P_i and P_o are given, respectively, by

$$f_{P_i}(\Theta_i) = \begin{cases} \frac{1}{2\pi}, & 0 \leq \Theta_i \leq 2\pi, \\ 0, & \text{otherwise}, \end{cases} \tag{4.3.1a}$$

$$f_{P_o}(\Theta_o) = \begin{cases} \frac{1}{2\pi}, & 0 \leq \Theta_o \leq 2\pi, \\ 0, & \text{otherwise}. \end{cases} \tag{4.3.1b}$$

Since the locations of P_i and P_o are independent from each other, their joint PDF is given by

$$f(\Theta_i, \Theta_o) = \begin{cases} \frac{1}{4\pi^2}, & 0 \leq \Theta_i, \Theta_o \leq 2\pi, \\ 0, & \text{otherwise}. \end{cases} \tag{4.3.2}$$

The probability that $\theta \leq \Theta$, which is also the cumulative distribution function (CDF) of θ, can be derived using the following integral [BHPC04]:

$$\begin{aligned} F(\Theta) &= P(\theta \leq \Theta) \\ &= \int\int_{\Omega} f(\theta_i, \theta_o) d\theta_o d\theta_i, \end{aligned} \tag{4.3.3}$$

where Ω is the space of locations of entry and exit points P_i and P_o such that $\theta \leq \Theta$ and $0 \leq \Theta \leq 2\pi$. $P(\theta \leq \Theta) = 0$ for $\Theta < 0$ and $P(\theta \leq \Theta) = 1$ for $\Theta > 2\pi$. From the observation of Figure 4.1 Equation (4.3.3) can be rewritten as

$$\begin{aligned} F(\Theta) &= P(\theta \leq \Theta) \\ &= \frac{1}{4\pi^2} \left(\int_0^{\Theta} \int_0^{\Theta+\theta_i} + \int_{\Theta}^{2\pi-\Theta} \int_{\theta_i-\Theta}^{\Theta+\theta_i} + \int_{2\pi-\Theta}^{2\pi} \int_{\theta_i-\Theta}^{2\pi} \right) d\theta_o d\theta_i \\ &= 4\pi\Theta - 4\Theta^2, \qquad 0 \leq \Theta \leq 2\pi. \end{aligned} \tag{4.3.4}$$

The PDF of θ can be derived by taking the derivative of Equation (4.3.4) and is given by

$$f(\Theta) = \begin{cases} \frac{1}{\pi}(1 - \frac{\Theta}{2\pi}), & 0 \leq \Theta \leq 2\pi, \\ 0. & \text{otherwise.} \end{cases} \tag{4.3.5}$$

The next step is to use the PDF of θ, and the expression of the traveling time t_{WLAN} as a function of θ to obtain the PDF of t_{WLAN}.

From the geometric configuration in Figure 4.1 and by using the cosine formula, the following equation is obtained:

$$(v t_{\mathrm{WLAN}})^2 = 2R^2(1 - \cos\theta). \tag{4.3.6}$$

Thus,

$$\begin{aligned} t_{\text{WLAN}} &= g(\theta) \\ &= \sqrt{\frac{2R^2}{v^2}(1-\cos\theta)}. \end{aligned} \tag{4.3.7}$$

Using the theorem stated in [Pap65, Equation 5.6], the PDF of t_{WLAN} is expressed as

$$f(T) = \sum_1 n \frac{f(\theta_n)}{|g'(\theta_n)|}, \tag{4.3.8}$$

where $\theta_1, ..., \theta_n$ are the roots of function $g(\theta)$, and $g'(\cdot)$ is the derivative of $g(\cdot)$.

In Equation (4.3.7), for $g(\theta)$ there are two roots, θ_1 and θ_2, which are expressed as

$$\theta_1 = \arccos\left(1 - \frac{v^2 t_{\text{WLAN}}^2}{2R^2}\right), \tag{4.3.9}$$

$$\theta_2 = 2\pi - \arccos\left(1 - \frac{v^2 t_{\text{WLAN}}^2}{2R^2}\right). \tag{4.3.10}$$

From (4.3.7), $g'(\theta)$ is expressed as

$$g'(\theta) = \frac{R\sin\theta}{v\sqrt{2(1-\cos\theta)}}. \tag{4.3.11}$$

So

$$
\begin{aligned}
|g'(\theta_1)| &= \left| \frac{R\sin\left(\arccos\left(1-\frac{v^2t_{\text{WLAN}}^2}{2R^2}\right)\right)}{v\sqrt{2\left[1-\cos\left(\arccos\left(1-\frac{v^2t_{\text{WLAN}}^2}{2R^2}\right)\right)\right]}} \right| \\
&= R\sqrt{1-\frac{v^2t_{\text{WLAN}}^2}{4R^2}}, \qquad (4.3.12) \\
|g'(\theta_2)| &= \left| \frac{R\sin\left(2\pi-\arccos\left(1-\frac{v^2t_{\text{WLAN}}^2}{2R^2}\right)\right)}{v\sqrt{2\left[1-\cos\left(2\pi-\arccos\left(1-\frac{v^2t_{\text{WLAN}}^2}{2R^2}\right)\right)\right]}} \right| \\
&= R\sqrt{1-\frac{v^2t^2}{4R^2}}, \qquad (4.3.13)
\end{aligned}
$$

and

$$
\begin{aligned}
f(\theta_1) &= \frac{1}{\pi}\left[1-\frac{\arccos\left(1-\frac{v^2t_{\text{WLAN}}^2}{2R^2}\right)}{2\pi}\right] \qquad (4.3.14) \\
f(\theta_2) &= \frac{1}{\pi}\left[1-\frac{2\pi-\arccos\left(1-\frac{v^2t_{\text{WLAN}}^2}{2R^2}\right)}{2\pi}\right]. \qquad (4.3.15)
\end{aligned}
$$

Thus, using Equations (4.3.8), (4.3.12) and (4.3.14) the PDF of t_{WLAN} is calculated by

$$
\begin{aligned}
f(T) &= \begin{cases} \frac{f(\theta_1)}{|g'(\theta_1)|}+\frac{f(\theta_2)}{|g'(\theta_2)|}, & 0 \le T \le \frac{2R}{v}, \\ 0, & \text{otherwise}, \end{cases} \\
&= \begin{cases} \frac{2}{\pi\sqrt{4R^2-v^2T^2}}, & 0 \le T \le \frac{2R}{v}, \\ 0, & \text{otherwise}. \end{cases} \qquad (4.3.16)
\end{aligned}
$$

The third step is to use the PDF of t_{WLAN} to obtain the CDF of t_{WLAN}, which is derived from the integral of Equation (4.3.16) as:

$$\begin{aligned} F(T) &= \Pr[t \leq T] \\ &= \int_0^T f(T)dT \\ &= \begin{cases} 1, \ , & \frac{2R}{v} < T, \\ \frac{2}{\pi} \arccos\left(\frac{vT}{2R}\right) \ , & 0 \leq T \leq \frac{2R}{v}. \end{cases} \end{aligned} \tag{4.3.17}$$

(4.3.18)

A time threshold parameter $T1$ is introduced to make handover decisions: whenever the estimated traveling time t_{WLAN} is greater than $T1$, the MT will initiate the handover procedure. A handover failure occurs when the traveling time inside the WLAN cell is shorter than the handover latency from the cellular network to the WLAN, τ_i. Thus, using Equation (4.3.17) the probability of a handover failure for the method using the threshold $T1$ is given by

$$P_f = \begin{cases} \frac{2}{\pi}\left[\arcsin\left(\frac{v\tau_i}{2R}\right) - \arcsin\left(\frac{vT1}{2R}\right)\right], & 0 \leq T1 \leq \tau_i, \\ 0, & \tau_i < T1. \end{cases} \tag{4.3.19}$$

By using (4.3.19), an equation which can be used by the MT to calculate the value of $T1$ for a particular value of P_f when $0 < P_f < 1$:

$$T1 = \frac{2R}{v} \sin\Big(\arcsin\left(\frac{v\tau_i}{2R}\right) - \frac{\pi}{2}P_f\Big). \tag{4.3.20}$$

To calculate $T1$, the speed of the MT v and the handover latency τ_i need to be obtained. In this research, the knowledge of v and τ_i is assumed, and they can be measured by using accelerometers [ZL07] and the technique described in [MA06], respectively.

4.3.2 Time Threshold Calculation for Minimizing Unnecessary Handovers

The purpose of the time threshold calculation presented in this section is to keep the number of unnecessary handovers under a desirable threshold. The method to calculate the time threshold is similar to the one used in Section 4.3.1, and it is explained below.

An unnecessary handover occurs if the traveling time inside the WLAN cell is shorter than the sum of the handover time into (τ_i) and out of (τ_o) the WLAN cell.

Similar to the arguments used in Section 4.3.1, another parameter $T2$ ($T1 < T2$) is introduced to minimize the probability of unnecessary handovers. By using (4.3.17) the probability of an unnecessary handover is calculated as

$$P_u = \begin{cases} \frac{2}{\pi}\left[\arcsin\left(\frac{v(\tau_i+\tau_o)}{2R}\right) - \arcsin\left(\frac{vT2}{2R}\right)\right], & 0 \leq T2 \leq (\tau_i + \tau_o), \\ 0, & (\tau_i + \tau_o) < T2. \end{cases} \tag{4.3.21}$$

Thus

$$T2 = \frac{2R}{v}\sin\left(\arcsin\left(\frac{v(\tau_i + \tau_o)}{2R}\right) - \frac{\pi P_u}{2}\right) \tag{4.3.22}$$

Equation (4.3.22) is derived from (4.3.21) for a particular value of P_u when $0 < P_u < 1$.

Parameters $T1$ and $T2$ depend on values of constants P_f and P_u which are selected by system designers. They also depend on measurement of v, R, τ_i and τ_o. The parameter $T2$ can be further adjusted dynamically to encourage or discourage handovers to WLAN by considering other performance criteria such as network load.

4.4 Conclusions

In this chapter, a method to estimate the handover necessity into a WLAN cell is presented. This method is based on two parts: traveling time estimation and time threshold calculation. The traveling time estimation relies on the RSS measurements and the speed of the MT. The time thresholds are calculated based on various network parameters such as tolerable handover failure probability or unnecessary handover probability, the radius of the WLAN cell and the handover latency. This method is able to reduce the number of handover failures and unnecessary handovers up to 80% and 70%, comparing with the conventional RSS threshold based [VRWF03] and hysteresis based [LLGD08]. Its performance is further discussed in Chapter 7.

Chapter 5

A New Method for Handover Target Selection

5.1 Introduction

In 4G environment, there are likely multiple candidate networks when the handover decision occurs. Thus, It is desirable to devise algorithms which select the most efficient network among all candidates as the handover target, so that users can benefit from the access in an "always best connected" mode [GAM05].

Some solutions to handover target selection algorithms were discussed in Section 2.4. In [ZM04, ZM06], the authors proposed a handover target selection method relying on a cost function which calculates the "cost" of possible target networks. Such cost is the sum of the cost of each QoS parameter, including the bandwidth, battery power and delay. The network with the minimum cost is selected as the handover target. This method increased the percentage of user satisfied requests and reduced the call blocking probability. However, the authors did not discuss how the QoS factors were normalized or how the weights for the QoS factors were assigned.

In this chapter, a handover target selection method is introduced. This method is devised for maximizing the user satisfaction level, by selecting the "best" network as the handover target among multiple candidate networks. It involves two algorithms: weights distribution and cost factor calculation.

This chapter is organized as follows: Section 5.2 presents the weights distribution algorithm, and Section 5.3 introduces the cost factor calculation algorithm.

5.2 Weights Distribution

The weights distribution (WD) algorithm takes user preferences and the power level of the MT as inputs, and generates weight factors for various handover decision parameters as outputs, as discussed in Section 3.3.

User preferences are user specified importance levels for the network parameters. These parameters are: available bandwidth, monetary cost and security. Four importance levels are defined: high, medium, low and none. The default level is low. Besides user preferences, the battery power level of the MT is also taken as an input for WD. WD calculates weight factors for available bandwidth, monetary cost, security and power consumption using a method described below.

Firstly, the following assumptions are made:

- The battery power level of the MT is p_w, where $0 < p_w \leq 1$, ($p_w = 0$ means the battery power runs out and $p_w = 1$ means the battery has the maximum power).
- The weight factors of the four network parameters, available bandwidth, monetary cost, security and power consumption, are w_B, w_M, w_S and w_P, respectively, where $w_P = 1 - p_w$ and $w_B + w_M + w_S + w_P = 1$.
- The factors of the importance levels of high, medium, low and none are

i_H, i_M, i_L and 0, respectively, where their values are decided by the mobile system designer, and $0 < i_H < i_M < i_L < 1$.

- The numbers of different importance levels the user has specified are n_H, n_M, n_L and n_N, respectively, where $n_H + n_M + n_L + n_N = 3$ (since the total number of the network parameters that a user could specify is three).
- The weight factors of the four importance levels, after adjusted to user preferences and battery power level, are w_{i_H}, w_{i_M}, w_{i_L} and w_{i_N}, respectively.

The objective is to calculate w_B, w_M, w_S and w_P based on the inputs of user preferences and battery power level. The following equations are obtained:

$$n_H \times w_{i_H} + n_M \times w_{i_M} + n_L \times w_{i_L} + n_N \times w_{i_N} = p_w \tag{5.2.1a}$$

$$w_{i_M} = w_{i_H} \times \frac{i_M}{i_H} \tag{5.2.1b}$$

$$w_{i_L} = w_{i_H} \times \frac{i_L}{i_H} \tag{5.2.1c}$$

$$w_{i_N} = 0 \tag{5.2.1d}$$

Substitute w_{i_M}, w_{i_L} and w_{i_N} in (5.2.1a) then

$$n_H \times w_{i_H} + n_M \times w_{i_H} \times \frac{i_M}{i_H} + n_L \times w_{i_H} \times \frac{i_L}{i_H} = p_w \tag{5.2.2}$$

So the weights of four importance levels are calculated by using the following equations

$$w_{i_H} = \frac{i_H p_w}{n_H i_H + n_M i_M + n_L i_L} \tag{5.2.3a}$$

$$w_{i_M} = \frac{i_M p_w}{n_H i_H + n_M i_M + n_L i_L} \tag{5.2.3b}$$

$$w_{i_L} = \frac{i_L p_w}{n_H i_H + n_M i_M + n_L i_L} \tag{5.2.3c}$$

$$w_{i_N} = 0 \qquad (5.2.3d)$$

Through Equation (5.2.3) the mobile system is able to assign weights for the four network parameters according to user preferences and battery power level. For example, the system designer specifies the factors of the four importance levels as 0.8 (i_H), 0.4 (i_M), 0.2 (i_L) and 0; the user assigns importance levels of available bandwidth, monetary cost and security to be high, medium and low, respectively (so that $n_H = n_M = n_L = 1$); the battery has 80% (p_w) of power left. Then it is calculated that the weights of available bandwidth, monetary cost, security and power consumption are 46% (w_B), 23% (w_M), 11% (w_S) and 20% (w_P).

5.3 Cost Factor Calculation

The cost factor calculation (CFC) algorithm evaluates the cost for making a handover to any candidate network by using a cost function. It takes various network parameters and their weights as inputs and generates cost factors for all candidate networks, as discussed in Section 3.3. The network with the lowest cost factor is selected as the handover target. The cost factor C_i, which provides a measurement of the cost of a certain network i, is calculated using the following function:

$$C_i = C\left(w_B B_i, w_M M_i, w_S S_i, w_P P_i\right) \qquad (5.3.1)$$

where $C(\cdot)$ is the cost function, B_i, M_i, S_i and P_i stand for available bandwidth (in Mbps), monetary cost per minute (in cents), security level (on a scale of 1 to 10, from very low to very high) and power consumption level (on a scale of 1 to 10, from very low to very high), and w_B, w_M, w_S and w_P are their weights obtained from the WD algorithm.

Since each network parameter has a different unit, it is necessary to normalize them in the cost function. The normalized cost factor for the number of n candidate networks is:

$$\begin{aligned} C_i \;=\; & \frac{w_B\left(1/B_i\right)}{\max\left(\left(1/B_1\right),\ldots,\left(1/B_n\right)\right)}+\frac{w_M M_i}{\max\left(M_1,\ldots,M_n\right)}+ \\ & \frac{w_S\left(1/S_i\right)}{\max\left(\left(1/S_1\right),\ldots,\left(1/S_n\right)\right)}+\frac{w_P P_i}{\max\left(P_1,\ldots,P_n\right)} \end{aligned} \tag{5.3.2}$$

Lastly, the network with a cost factor of $\min\left(C_1,\ldots,C_n\right)$ is selected as the handover target.

5.4 Conclusions

In this chapter, a method to select the handover target is presented. This method is based on two parts: weights distribution and cost factor calculation. Weights of various network parameters are generated based on user preferences and the power level of the MT, and cost factors of candidate networks are calculated using a cost function. The network with the lowest cost factor is selected as the handover target. This method is able to maximize users' satisfaction up to 50%, comparing with methods that consistently choose one access network. Its performance is further discussed in Chapter 7.

Chapter 6

A New Method for Handover Triggering Condition Estimation

6.1 Introduction

It is expected that smart phones equipped with intelligent algorithms will seek to maximize the use of low-cost and high-bandwidth WLAN connections, whenever available, as an alternative to cellular access. In conjunction with this, it is important to minimize connection breakdowns as handovers occur between these access networks, so that session continuity is maintained without perceivable interruptions [DDF+07]. Thus, the handover triggering point becomes a critical issue as the user travels across the WLAN cell coverage. A handover should neither be triggered too late resulting in a connection breakdown, nor too early resulting in the wastage of available WLAN resource. There is a tradeoff between connection breakdown probability and WLAN usage, and both of them are related to the speed of the MT.

Methods of determining the handover triggering point have been proposed in a number of earlier studies [LLGD08, PYK+03, YPMM01]. However, to date there is no method that provides the user with a mechanism for managing the tradeoff between connection breakdowns and WLAN usage while dynamically adapting to the speed of the MT.

In this chapter, a handover triggering condition estimation (HTCE) method is introduced. This method attempts to estimate the optimal handover triggering point when a MT needs to initiate a handover back from a WLAN to a cellular network. HTCE is devised to keep the connection breakdown probability below user adjustable limits under different MT speeds, as well as to provide the user with control over the tradeoff between connection breakdown probability and WLAN usage.

The rest of this chapter is organized as follows. Section 6.2 provides an overview of the HTCE process, and Section 6.3 describes the RSS threshold calculation algorithm.

6.2 Overview of the Handover Triggering Condition Estimation Process

HTCE determines a proper time to initiate a handover out of the currently connected network to prevent connection breakdowns as well as to maximize the usage of the preferred network. It takes AP power level, RSS samples, the estimated radius of the WLAN (the radius of the WLAN cell using RSS samples and the AP transmit power, as the described in Chapter 4), the velocity of the MT, the handover latency and the connection breakdown probability requirement as inputs, generating the handover triggering condition as its output. The block diagram of HTCE is shown in Figure 3.6 in Chapter 3.

Figure 6.1 shows the trajectory of a MT traveling over an area over which cellular network service is available and is also partially covered with a WLAN cell. The

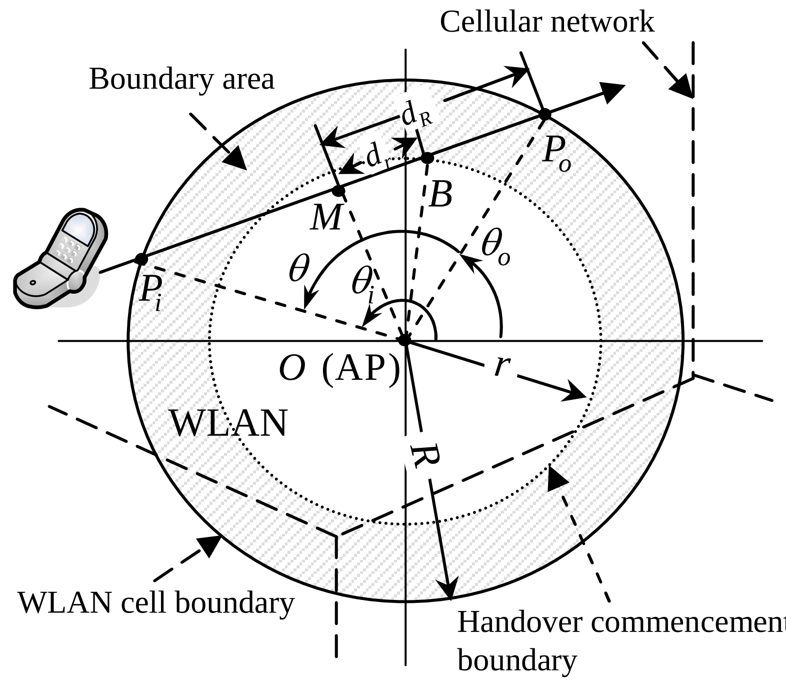

Figure 6.1: Scheme diagram of WLAN to cellular network handover triggering condition estimation mechanism.

MT enters and exits the WLAN cell at points P_i and P_o, respectively, following a straight line. M is the middle point of the section of the trajectory inside the WLAN cell. The donut shaped (dashed) area is called "boundary area". The radii of the outer and inner circles enclosing the boundary area are R and r, and d_R and d_r represent the half length of the trajectory segments inside the outer and inner circles respectively. B is the intersection point of the trajectory and the inner circle.

When the MT enters the area of WLAN coverage at point P_i, the algorithm proposed in Chapter 4 is applied to determine whether a handover to the WLAN is beneficial or not. If a handover is necessary, the MT switches to the WLAN. Then, the MT continues its trajectory, and beyond the point M, the RSS of the AP starts deteriorating. Here, HTCE is used to determine the point at which the handover back to the cellular network should occur.

Overall operation of the proposed algorithm can be summarized as follows:

1. MT takes periodic measurements of WLAN RSS and calculates its moving average ($\overline{\text{RSS}}$);

2. Runs HTCE algorithm to estimate r and RSS at point B (RSS_B, discussed in Section 6.3);

3. Triggers a handover when $\overline{\text{RSS}}$ drops below RSS_B.

So, the handover triggering point is defined by RSS_B. In the following section, mathematical arguments for estimating the value of RSS_B is presented.

6.3 RSS Threshold Calculation

The purpose of the RSS threshold calculation in this section is to keep the number of connection breakdowns under a desirable threshold. For example, if the system designer has a requirement of limiting the probability of connection breakdowns under 10%, then the RSS threshold is adjusted to make the ratio of the number of connection breakdowns to the total number of handovers below 10%. The method to achieve this goal is using a handover commencement boundary, as shown in Figure 6.1. A handover out of the WLAN is triggered once the MT reaches the boundary at point B, i.e. the RSS from the WLAN drops below a threshold RSS_B. The mathematical arguments used to estimate the value of RSS_B are presented below.

Same as in Chapter 4, it is assumed that for a randomly selected trajectory, P_i and P_o are uniformly distributed over the WLAN cell boundary (Figure 6.1). The angles θ_i and θ_o are both uniformly distributed in $[0, 2\pi]$, and $\theta = |\theta_i - \theta_o|$. The PDF of θ is obtained by using Equation (4.3.5).

To find the handover commencement boundary, it is needed to determine the radius r of the inner circle. This radius is a dynamic parameter which depends on the handover delay (technology and protocol dependent), a desirable upper limit on handover breakdown probability, the speed of the MT and the averaged estimation error ratio for handover commencement distance (to compensate for shadowing effects).

In the following paragraphs, the algorithm to determine the value of r and RSS_B is described. Firstly, from the geometry configuration in Figure 6.1 and by using the cosine formula, the following equations are obtained:

$$(2d_R)^2 = R^2 + R^2 - 2R^2 \cos\theta, \tag{6.3.1a}$$

$$l_{OM}^2 = R^2 - {d_R}^2, \tag{6.3.1b}$$

$$d_r^2 = r^2 - l_{OM}^2. \tag{6.3.1c}$$

By substituting Equation (6.3.1b) in Equation (6.3.1c) and using algebraic manipulation, the following equations are obtained:

$$d_R = \sqrt{\frac{1}{2}R^2\,(1 - \cos\theta)}, \tag{6.3.2a}$$

$$d_r = \sqrt{r^2 - \frac{1}{2}R^2\,(1 + \cos\theta)}. \tag{6.3.2b}$$

A handover is triggered if one of the two conditions are satisfied: $C1$, the RSS from the WLAN drops below the RSS threshold RSS_B when the MT is traveling away from the WLAN AP; $C2$, the RSS from the WLAN is below RSS_B (and has never been above RSS_B) and starts continuously deteriorating, which means the MT travels past the middle point M. Thus, the remaining traveling time t_b inside the boundary area, i.e. the time spent on traveling from point B to P_o for $C1$ and from point M to P_o for $C2$, is determined as

$$t_b = \begin{cases} \frac{1}{v}\left(d_R - d_r\right), & C1, \\ \frac{1}{v}d_R, & C2, \end{cases} \tag{6.3.3}$$

By substituting Equation (6.3.2) in Equation (6.3.3), t_b is expressed as

$$t_b = \begin{cases} \frac{\sqrt{\frac{1}{2}R^2(1-\cos\theta)}-\sqrt{r^2-\frac{1}{2}R^2(1+\cos\theta)}}{v}, & C_1, \\ \frac{\sqrt{\frac{1}{2}R^2(1-\cos\theta)}}{v}, & C_2, \end{cases}$$

where $t_b \in \left[0, \frac{\sqrt{R^2-r^2}}{v}\right]$, with the maximum value at $\theta = 2\cos^{-1}\left(\frac{r}{R}\right)$ and $\theta = 2\pi - 2\cos^{-1}\left(\frac{r}{R}\right)$.

By using the definition of CDF [Pap65], the CDF of t_b, $F(T)$ is expressed as,

$$\begin{aligned} F(T) &= \Pr[t_b \leq T] \\ &= \begin{cases} p, , & 0 \leq T \leq \frac{1}{v}\sqrt{R^2 - r^2}, \\ 1\,, & \text{otherwise.} \end{cases} \end{aligned} \tag{6.3.4}$$

where

$$p = \begin{cases} p_1 &= \Pr\left[\frac{d_R - d_r}{v} \leq T\right], \quad C_1 \\ p_2 &= \Pr\left[\frac{d_R}{v} \leq T\right], \quad C_2. \end{cases} \tag{6.3.5}$$

The probability of $\frac{d_R - d_r}{v} \leq T$, p_1, is calculated as

$$
\begin{aligned}
p_1 &= \Pr\left[\frac{\sqrt{\frac{1}{2}R^2\left(1-\cos\theta\right)} - \sqrt{r^2 - \frac{1}{2}R^2\left(1+\cos\theta\right)}}{v} \leq T\right] \\
&= \Pr\left[\frac{\sqrt{\frac{1}{2}R^2\left(1-\cos\theta\right)} - \sqrt{r^2 - \frac{1}{2}R^2\left(1+\cos\theta\right)}}{v} \leq T\right] \\
&= \Pr\left[\cos\theta \leq \frac{2r^2}{R^2} - \frac{(R^2 - r^2 - v^2T^2)^2}{2v^2T^2R^2} - 1\right] \\
&= \Pr[a_T \leq \theta \leq 2\pi - a_T]\,,
\end{aligned}
\tag{6.3.6}
$$

where

$$
a_T = \cos^{-1}\!\left(\frac{2r^2}{R^2} - \frac{(R^2 - r^2 - v^2T^2)^2}{2v^2T^2R^2} - 1\right). \tag{6.3.7}
$$

Then, by integrating $f(\Theta)$ given in Equation (4.3.5), p_1 is obtained as

$$
\begin{aligned}
p_1 &= \int_{a_T}^{2\pi - a_T} f(\Theta)d\Theta \\
&= 1 - \frac{1}{\pi}\cos^{-1}\!\left(\frac{2r^2}{R^2} - \frac{(R^2 - r^2 - v^2T^2)^2}{2v^2T^2R^2} - 1\right),
\end{aligned}
\tag{6.3.8}
$$

which is the CDF of θ within the range stated in Equation (6.3.6).

The probability of $\frac{d_R}{v} \leq T$, p_2, is calculated as

$$
\begin{aligned}
p_2 &= \Pr\left[\frac{\sqrt{\frac{1}{2}R^2\left(1-\cos\theta\right)}}{v} \leq T\right] \\
&= \Pr\left[\cos\theta \geq 1 - \frac{2v^2T^2}{R^2}\right] \\
&= \Pr\left[0 \leq \theta \leq \cos^{-1}\!\left(1 - \frac{2v^2T^2}{R^2}\right) \;\cup\; 2\pi - \cos^{-1}\!\left(1 - \frac{2v^2T^2}{R^2}\right) \leq \theta \leq 2\pi\right].
\end{aligned}
\tag{6.3.9}
$$

Then, by integrating $f(\Theta)$ given in Equation (4.3.5), p_2 is obtained as

$$
\begin{aligned}
p_2 &= \int_0^{\cos^{-1}\left(1-\frac{2v^2T^2}{R^2}\right)} f(\Theta)d\Theta + \int_{2\pi-\cos^{-1}\left(1-\frac{2v^2T^2}{R^2}\right)}^{2\pi} f(\Theta)d\Theta \\
&= \frac{1}{\pi}\cos^{-1}\left(1-\frac{2v^2T^2}{R^2}\right). \qquad (6.3.10)
\end{aligned}
$$

which is the CDF of θ within the range stated in Equation (6.3.9).

So the ultimate expression of the CDF of t_b, $F(T)$, is:

$$
F(T) = \Pr[t_b \le T] = \begin{cases} 1 - \frac{1}{\pi}\cos^{-1}\left(\frac{2r^2}{R^2} - \frac{(R^2-r^2-v^2T^2)^2}{2v^2T^2R^2} - 1\right), & 0 \le T \le \frac{1}{v}\sqrt{R^2-r^2} \ \cap \ C_1, \\ \frac{1}{\pi}\cos^{-1}\left(1-\frac{2v^2T^2}{R^2}\right), & 0 \le T \le \frac{1}{v}\sqrt{R^2-r^2} \ \cap \ C_2, \\ 1, & \text{otherwise}. \end{cases} \qquad (6.3.11)
$$

A connection breakdown occurs when the traveling time inside the boundary area is less than the handover delay from the WLAN to the cellular network, τ_o. The probability of a connection breakdown P_b is calculated as,

$$
P_b = G(r) = \begin{cases} 1, & R < r, \\ 1 - \frac{1}{\pi}\cos^{-1}\left(\frac{2r^2}{R^2} - \frac{(R^2-r^2-v^2\tau_o^2)^2}{2v^2\tau_o^2R^2} - 1\right), & R - v\tau_o \le r \le R \ \cap \ C_1, \\ 0, & r < R - v\tau_o \ \cap \ C_1, \\ \frac{1}{\pi}\cos^{-1}\left(1-\frac{2v^2\tau_o^2}{R^2}\right), & C_2. \end{cases} \qquad (6.3.12)
$$

By using (6.3.12), an equation which can be used by the MT to calculate the value of r for a particular value of P_b when $0 < P_b < 1$:

$$
\begin{aligned}
r &= G^{-1}(P_b) - C_aR \\
&= \sqrt{v^2\tau_o^2 + R^2 - v\tau_oR\sqrt{2\left[1-\cos\left(\pi - \pi P_b\right)\right]}} - C_aR, \qquad (6.3.13)
\end{aligned}
$$

where $G^{-1}(P_b)$ is the inverse function of $G(r)$, and C_a is a channel adjustment parameter. RSS measurements are widely used in distance estimation techniques because they require no additional hardware, however, shadowing degrades the accuracy of estimation significantly [HAK+09]. HTCE applies the RSS moving average method [Sin07] to compensate for shadowing effects. In our earlier study [YŞM08], we observed that at lower velocities the estimation based on the RSS introduced a deviation of 10% on average from the real distance (caused by shadowing effects). In order to accommodate this fluctuation, conservatively, we include an empirical compensation factor (C_a) based on the velocity of the MT. C_a is calculated as

$$C_a = \begin{cases} \frac{10\%}{v}, & \text{if } v > 3.6\ \mathrm{km/h}, \\ 10\%, & \text{else.} \end{cases} \tag{6.3.14}$$

Finally, by using the log-distance path loss model [Rap02, Eq. 3.69a], RSS_B is obtained

$$\mathrm{RSS}_B = P_{\mathrm{Tx}} - \mathrm{PL}_{\mathrm{ref}} - 10\beta \log_{10} \frac{r}{d_{\mathrm{ref}}} + X_\sigma, \tag{6.3.15}$$

where P_{Tx} is the transmit power of the WLAN AP in dBm, d_{ref} is the distance between the AP and a reference point, $\mathrm{PL}_{\mathrm{ref}}$ is the path loss at the reference point in dB, β is the path loss exponent, and X_σ is a Gaussian distributed random variable with a mean of zero and a standard deviation σ in dB.

6.4 Conclusions

In this chapter, a method for estimating the handover triggering condition has been presented. This method calculates a RSS threshold for triggering a handover based on the estimated radius of the WLAN cell, handover latency, speed of the

MT and connection breakdown tolerance. HTCE is able to provide the user with control over the tradeoff between connection breakdowns and WLAN usage. Its performance is further discussed in Chapter 7.

Chapter 7

Results and Discussions

7.1 Introduction

In this chapter, results of the proposed methods are presented. The proposed methods are compared against two other methods: (a) the fixed RSS threshold based method [VRWF03], in which handovers between the cellular network and the WLAN are initiated when the RSS from the WLAN reaches a fixed threshold, and (b) the hysteresis based method [LLGD08], in which a hysteresis is introduced to prevent the ping-pong effect [YPMM01].

The results are divided into three parts. In Sections 7.2, 7.4 and 7.3, results of HNE, HTCE and HTS are provided, respectively. Both theoretical and simulation results are presented in these sections.

7.2 Results of Handover Necessity Estimation

In this section, the performance of the HNE method is demonstrated. Firstly the theoretical performance is analyzed in Section 7.2.1, and then the simulation results are provided in Section 7.2.2.

PARAMETER	SYMBOL	VALUE
WLAN radius	R	150 m [LLGD08]
AP transmit power	P_{Tx}	20 dBm [KUKR05]
Distance between the AP and the reference point	d_{ref}	1 m
Path loss at the reference point	PL_{ref}	40 dB [ZPK03]
Path loss exponent	β	3.5
Standard deviation of shadow fading	σ	4.3 dB
Handover delay from cellular network to WLAN	τ_i	2 s
Handover delay from WLAN to cellular network	τ_o	2 s
Tolerable handover failure probability	P_f	0.02
Tolerable unnecessary probability	P_u	0.04

Table 7.1: Parameters used in the HNE performance evaluation.

The parameters used in theoretical analysis and simulations of HNE are listed in Table 7.1.

7.2.1 Theoretical Analysis of HNE

In the fixed RSS threshold based method [VRWF03], a handover to the WLAN is triggered when the RSS from the WLAN is above a threshold, $\text{RSS1}_{\text{fixed}}$. Using Equation (4.3.17), the handover failure probability for the fixed RSS threshold based method is given by

$$P_{f_{\text{fixed}}} = \begin{cases} 1, & v\tau_i > 2R1_{\text{fixed}}, \\ \frac{2}{\pi}\sin^{-1}\left(\frac{v\tau_i}{2R1_{\text{fixed}}}\right), & 0 \le v\tau_i \le 2R1_{\text{fixed}}, \end{cases} \tag{7.2.1}$$

where $R1_{\text{fixed}}$ is the distance between the MT location and the AP of the WLAN cell when a handover into the WLAN occurs in the fixed RSS threshold based method. It is calculated by

$$R1_{\text{fixed}} = 10^{\frac{E_t - \text{RSS1}_{\text{fixed}}}{10\beta}}. \tag{7.2.2}$$

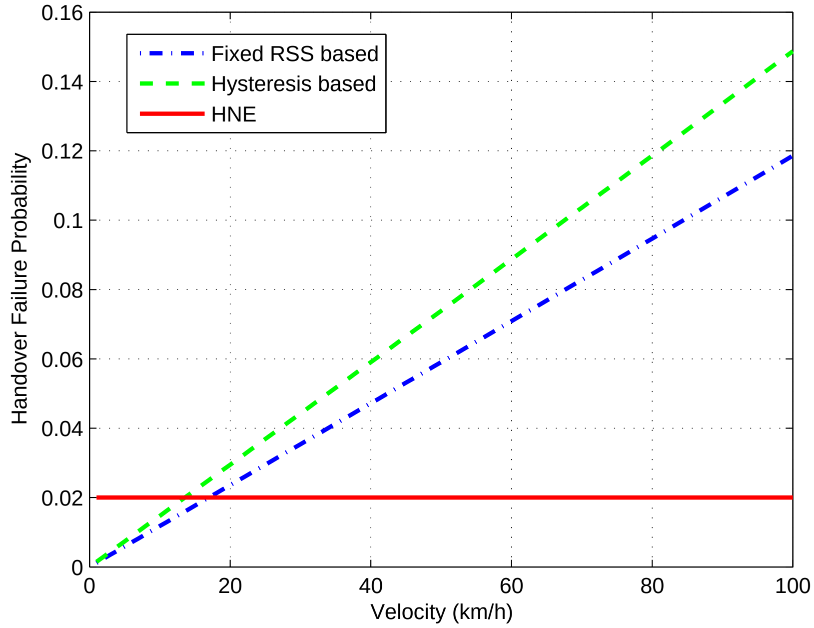

Figure 7.1: Probability of handover failures of the RSS threshold based [VRWF03], hysteresis based [LLGD08] and HNE methods.

In the hysteresis based method [LLGD08], a handover to the WLAN is triggered when the RSS from the WLAN is above a threshold plus a hysteresis, $\text{RSS1}_{\text{hyst}}+h_y$, where $\text{RSS1}_{\text{hyst}}$ is the RSS threshold and h_y is a constant representing the hysteresis.

Using Equation (4.3.17), the handover failure probability for the hysteresis based method is given by

$$P_{f_{\text{hyst}}} = \begin{cases} 1, & v\tau_i > 2R1_{\text{hyst}}, \\ \frac{2}{\pi}\sin^{-1}\left(\frac{v\tau_i}{2R1_{\text{hyst}}}\right), & 0 \leq v\tau_i \leq 2R1_{\text{hyst}}, \end{cases} \tag{7.2.3}$$

where $R1_{\text{hyst}}$ is the distance between the MT location and the AP of the WLAN cell when a handover into the WLAN occurs in the hysteresis based method. It is

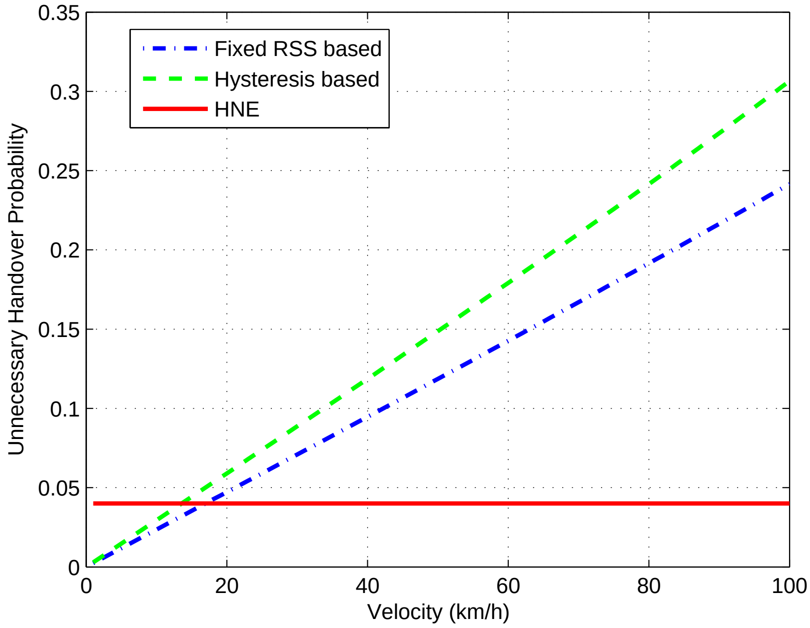

Figure 7.2: Probability of unnecessary handovers of the RSS threshold based [VRWF03], hysteresis based [LLGD08] and HNE methods.

calculated by

$$R1_{\text{hyst}} = 10^{\frac{E_t - \text{RSS1}_{\text{hyst}} + h_y}{10\beta}}. \tag{7.2.4}$$

Using Equation (4.3.17), the unnecessary handover probability for the fixed RSS threshold based method is given by

$$P_{u_{\text{fixed}}} = \begin{cases} 1, & v(\tau_i + \tau_o) > 2R1_{\text{fixed}}, \\ \frac{2}{\pi}\sin^{-1}\left(\frac{v(\tau_i+\tau_o)}{2R1_{\text{fixed}}}\right), & 0 \le v(\tau_i + \tau_o) \le 2R1_{\text{fixed}}. \end{cases} \tag{7.2.5}$$

The unnecessary handover probability for the hysteresis based method is given by

$$P_{u_{\text{hyst}}} = \begin{cases} 1, & v(\tau_i + \tau_o) > 2R1_{\text{hyst}}, \\ \frac{2}{\pi}\sin^{-1}\left(\frac{v(\tau_i+\tau_o)}{2R1_{\text{hyst}}}\right), & 0 \le v(\tau_i + \tau_o) \le 2R1_{\text{hyst}}. \end{cases} \tag{7.2.6}$$

The probabilities of handover failures and unnecessary handovers of the RSS threshold based ($R1_{\text{fixed}}$ = 150 m), hysteresis based ($R1_{\text{hyst}}$ = 120 m) and HNE methods are shown in Figures 7.1 and 7.2. Since HNE is designed to keep the probability of handover failures or unnecessary handovers below preset levels, even though the velocity of the MT increases, the probabilities remain the same. As illustrated by the figures, for higher velocities, HNE yields lower probability of handover failures and unnecessary handovers than the other two methods. Otherwise, for velocities less than 20 km/h, the other two methods yield marginally better results.

7.2.2 Simulation Results of HNE

MATLAB [MAT92] was used for the experiments, which generated 10000 random MT trajectories across the WLAN cell coverage area for speeds from 1 km/h to 100 km/h in 2 km/h increments. For each trajectory, a random WLAN cell entry point was chosen, and a uniformly distributed random angle between 0 and 2π was generated representing the movement direction of the MT.

Figures 7.3 and 7.4 denote the number of handover failures and unnecessary handovers of the RSS threshold based, hysteresis based and HNE methods under different velocities of the MT. Figure 7.5 shows the total number of handovers while applying these methods.

From the figures it can be seen that, with HNE, handover failures and unnecessary handovers are kept under the numbers of 200 and 500, respectively. The total number of handovers declines with the increasing velocity of the MT. In the RSS threshold and hysteresis based methods, the numbers of handover failures and unnecessary handovers increase sharply as the velocity increases. HNE is able to

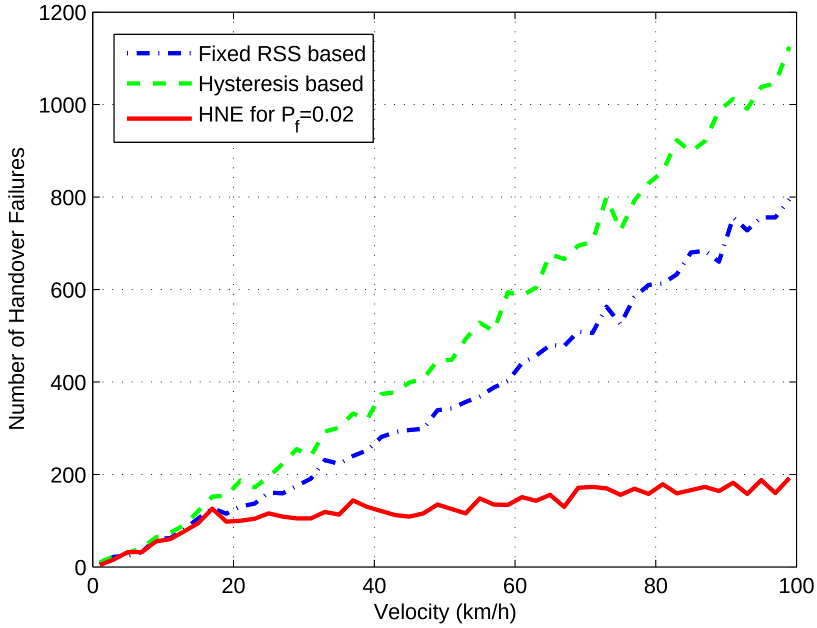

Figure 7.3: Number of handover failures of the RSS threshold based [VRWF03], hysteresis based [LLGD08] and HNE methods.

reduce the number of handover failures and unnecessary handovers up to 80%, when the velocity of the MT is up to 100 km/h.

For a better observation of the performance comparison, the ratios of the number of handover failures and unnecessary handovers to the total number of handovers are depicted in Figures 7.6 and 7.7, respectively. As can be seen, in HNE the ratio of the number of handover failures to the total number of handovers can be kept around the tolerable value of 0.02, and the ratio of the number of unnecessary handovers to the total number of handovers can be kept around the tolerable value of 0.04. HNE yields much better performance than the other two methods.

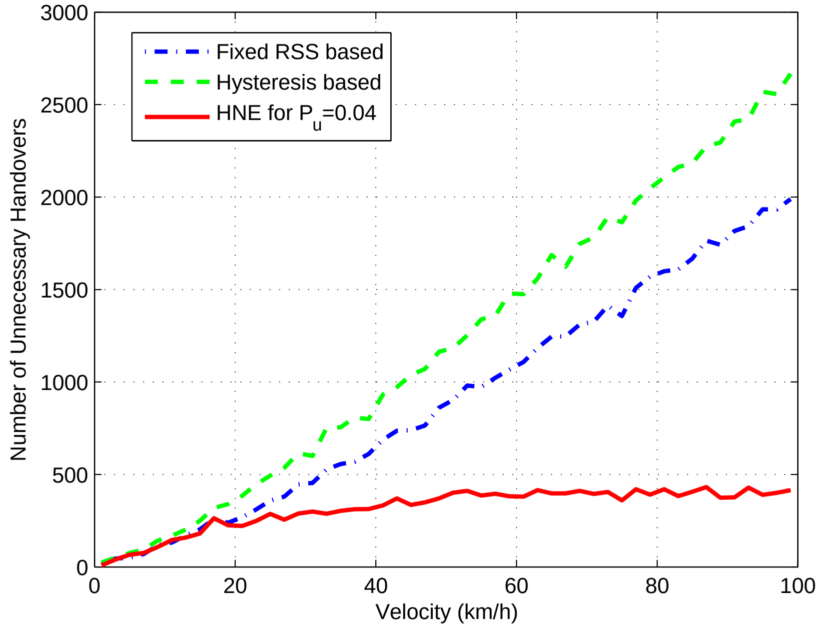

Figure 7.4: Number of unnecessary handovers of the RSS threshold based [VRWF03], hysteresis based [LLGD08] and HNE methods.

Network ID	1	2	3
Type	WLAN	WLAN	UMTS
Bandwidth (data rate)	2 Mbps	1 Mbps	384 Kbps
Monetary cost	3 cent/minute	2 cent/minute	5 cent/minute
Security level	1	2	7
Power consumption level	3	2	1

Table 7.2: Network parameters in the HTS performance evaluation.

7.3 Results of Handover Target Selection

In the performance evaluation for HTS, it is assumed that three candidate networks are available, including two WLANs and a UMTS. Their parameters are listed in Table 7.2.

MATLAB is used to generate 1000 sets of random user preferences, importance levels of bandwidth, monetary cost and security. A performance parameter called user's satisfaction is adopted to compare the results of consistently choosing one access network with using HTS. The user's satisfaction is calculated by adjusting

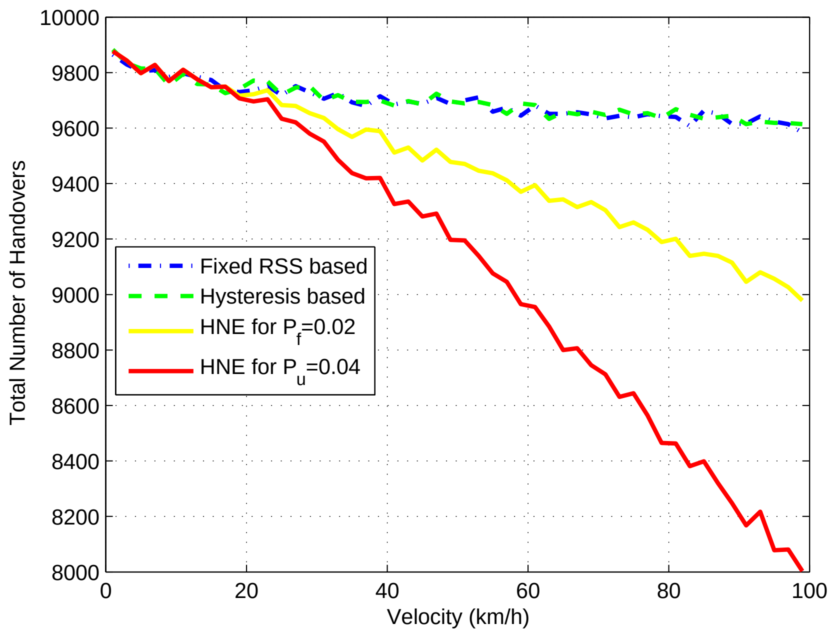

Figure 7.5: Number of handovers of the RSS threshold based [VRWF03], hysteresis based [LLGD08] and HNE methods.

the method in [HNH06] to add in more network parameters and their weights. The function of the user's satisfaction is

$$\text{User's Satisfaction} = \frac{Preferred\_Bandwidth - Actual\_Bandwidth}{Actual\_Bandwidth} \times w_B + \frac{Preferred\_Monetary\_Cost - Actual\_Monetary\_Cost}{Actual\_Monetary\_Cost} \times w_M + \frac{Preferred\_Security\_Level - Actual\_Security\_Level}{Actual\_Security\_Level} \times w_S, \tag{7.3.1}$$

where $Preferred\_Bandwidth$, $Preferred\_Monetary\_Cost$ and $Preferred\_Security\_Level$ are the preferred values of the network parameters specified by the user, $Actual\_Bandwidth$, $Actual\_Monetary\_Cost$ and $Actual\_Security\_Level$ are the actual values of the network parameters obtained in the simulation, and w_B, w_M and w_S are weights of these parameters calculated by the MT according to user preferences.

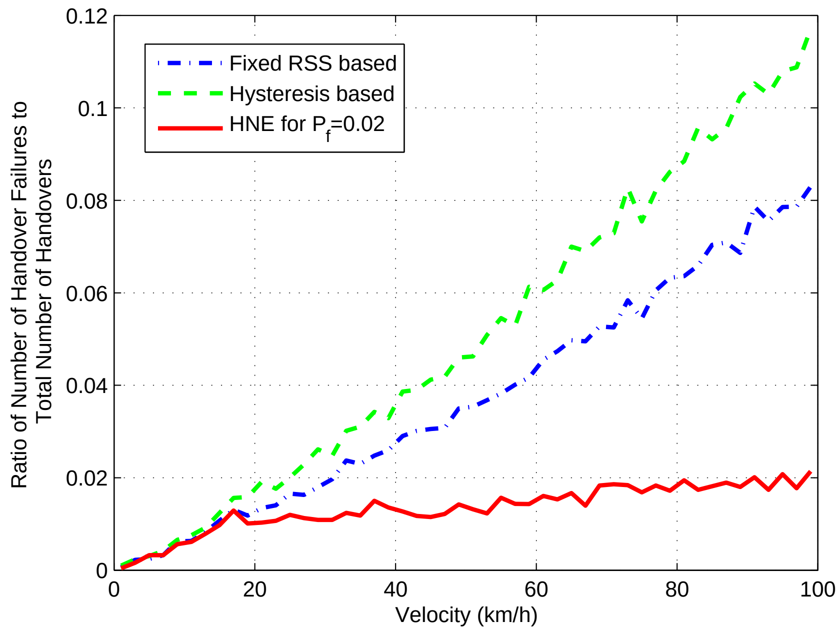

Figure 7.6: Ratio of the number of handover failures to the total number of handovers.

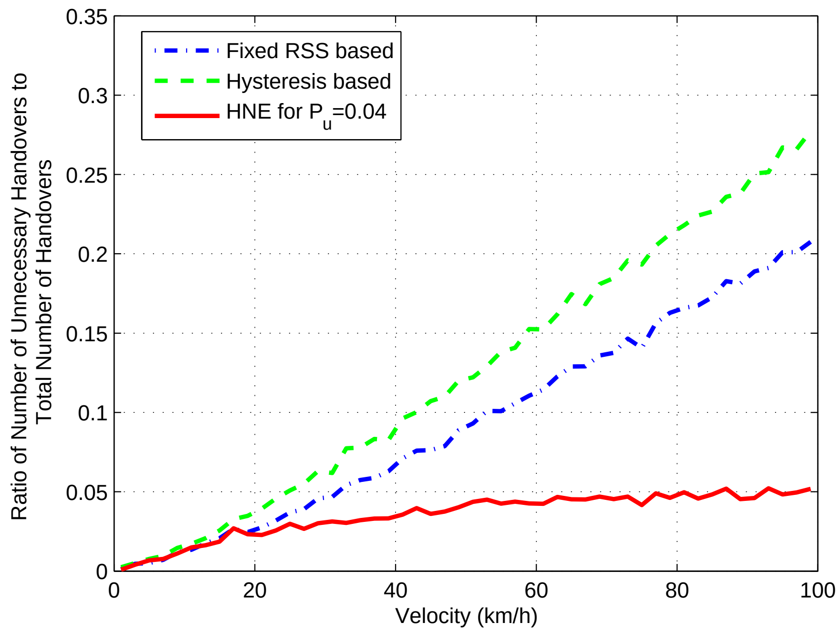

Figure 7.7: Ratio of the number of unnecessary handovers to the total number of handovers.

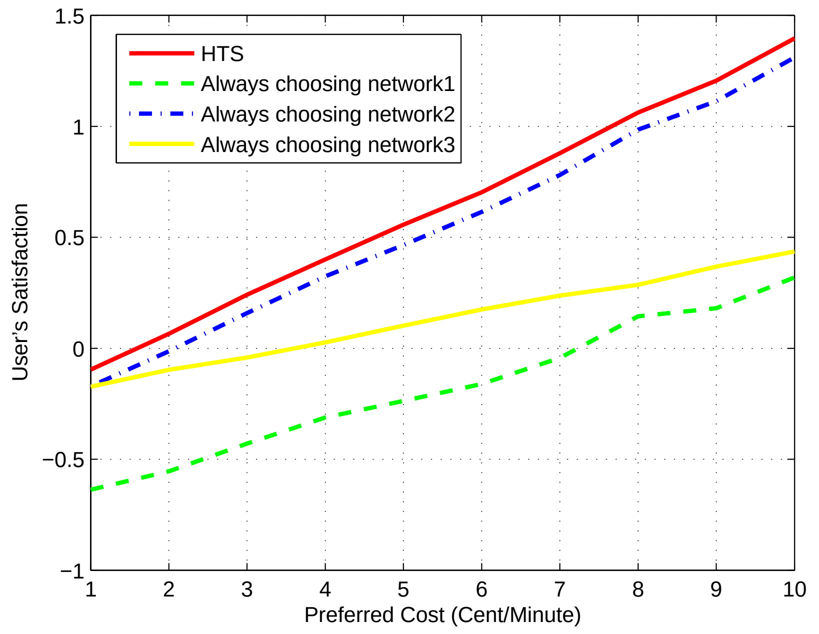

Figure 7.8: User's satisfaction based on different preferred monetary costs.

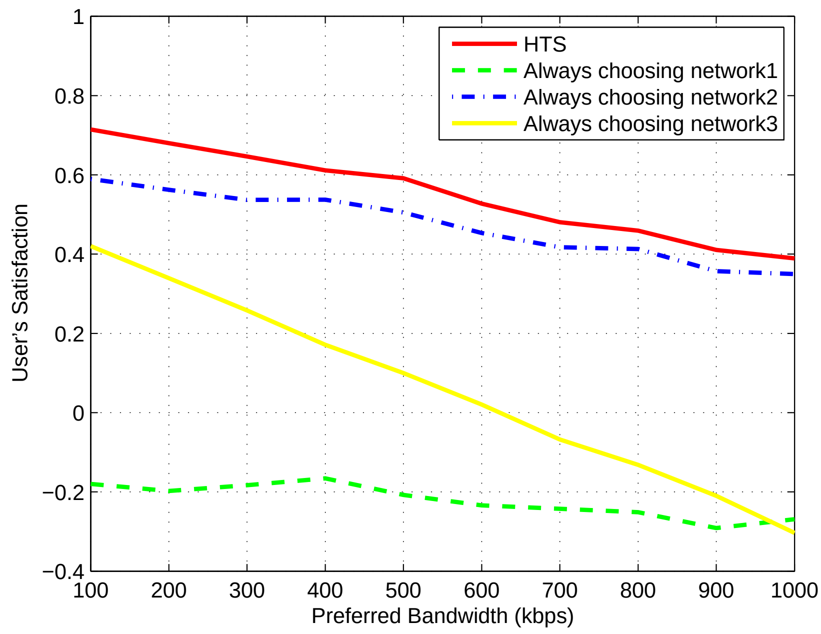

Figure 7.9: User's satisfaction based on different preferred bandwidth.

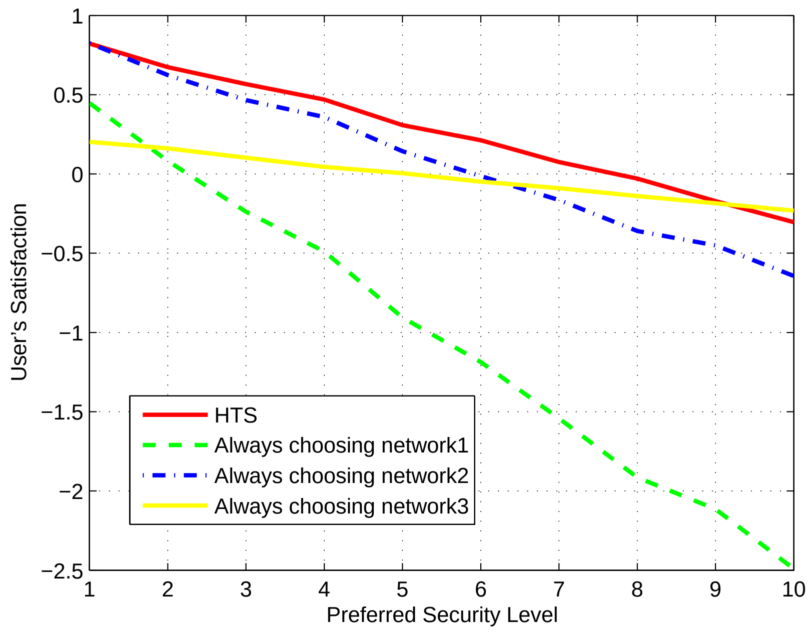

Figure 7.10: User's satisfaction based on different preferred security levels.

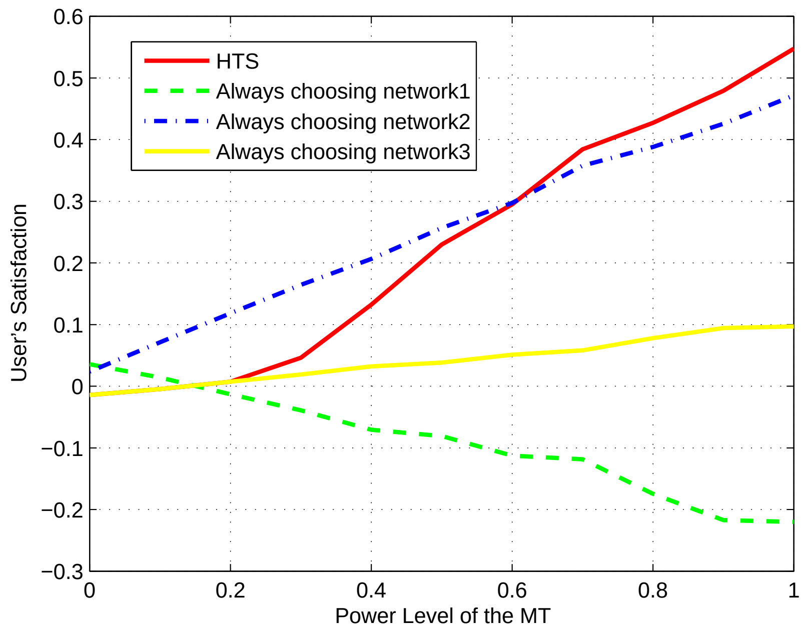

Figure 7.11: User's satisfaction based on different power levels of the MT.

Figures 7.8, 7.9, 7.10 and 7.11 show the user's satisfaction generated by adopting HTS and consistently choosing Network 1, Network 2 and Network 3.

In Figure 7.8, the power level of the MT p_w is set to be 100%, and $Preferred\_Bandwidth$ and $Preferred\_Security\_Level$ are set to be fixed values of 500 Kbps and 3, respectively. The average user's satisfaction is calculated based on different values of $Preferred\_Monetary\_Cost$. As can be seen, the highest user satisfaction is achieved by HTS no matter what budget is specified by the user. Similar results are observed in Figures 7.9 and 7.10 where different values of $Preferred\_Bandwidth$ and $Preferred\_Security\_Level$ are applied. Overall, the results indicate that HTS increases users' satisfaction and helps manage their budget or other requirements.

In Figure 7.11, the influence of different power levels of the MT on the user's satisfaction is simulated. It is shown that for power levels lower than 0.6, HTS achieves lower user's satisfaction than consistently choosing Network 2. This is because HTS tends to select Network 3 with low power consumption level when the power level of the MT is relatively low in order to extend the battery life, and the power consumption factor is not included when calculating user's satisfaction.

7.4 Results of Handover Triggering Condition Estimation

In this section, the performance of the HTCE method is demonstrated. Firstly the theoretical performance is analyzed in Section 7.4.1, and then the simulation results are provided in Section 7.4.2.

The parameters used in theoretical analysis and simulations of HTCE are listed in Table 7.3.

Parameter	Symbol	value
WLAN radius	R	150 m [LLGD08]
AP transmit power	P_{Tx}	20 dBm [KUKR05]
Distance between the AP and the reference point	d_{ref}	1 m
Path loss at the reference point	PL_{ref}	40 dB [ZPK03]
Path loss exponent	β	3.5 [Rap02, Table 3.2]
Standard deviation of shadow fading	σ	4.3 dB [Gud91]
Handover delay from cellular network to WLAN	τ_i	2 s [LLGD08]
Handover delay from WLAN to cellular network	τ_o	2 s [LLGD08]

Table 7.3: Parameters used in the HTCE performance evaluation.

7.4.1 Theoretical Analysis of HTCE

where $R2_{\text{hyst}}$ is the distance between the MT location and the AP of the WLAN cell when a handover out of the WLAN occurs in the hysteresis based method. It is calculated by

$$R2_{\text{hyst}} = 10^{\frac{E_t - \text{RSS2}_{\text{hyst}} + h_y}{10\beta}}. \qquad (7.4.1)$$

In the fixed RSS threshold based method [VRWF03], a handover out of the WLAN is triggered when the RSS from the WLAN is below a threshold, $\text{RSS2}_{\text{fixed}}$. Using

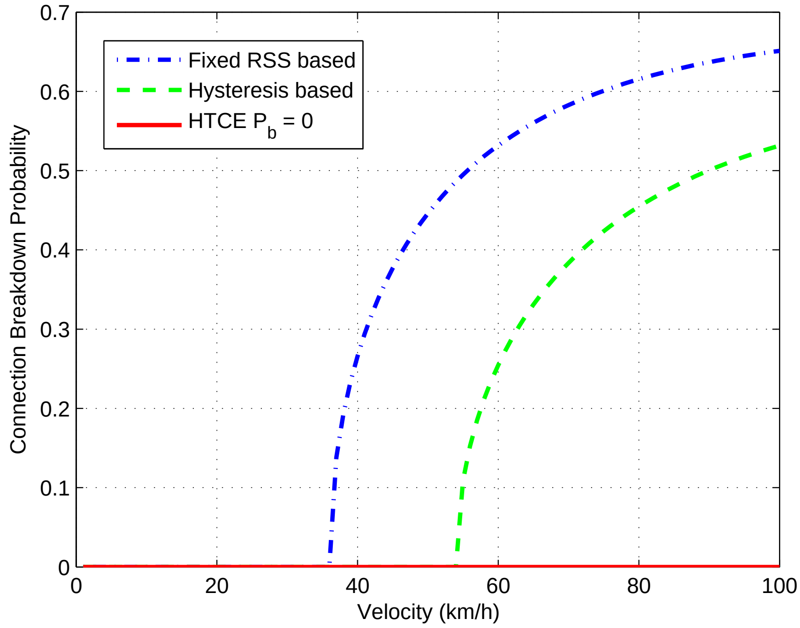

Figure 7.12: Probability of connection breakdowns of the fixed RSS threshold based [VRWF03], hysteresis based [LLGD08] and HTCE methods.

Equation (6.3.12), the connection breakdown probability for the fixed RSS threshold based method is given by

$$P_{b_{\text{fixed}}} = \begin{cases} 1, & R < R2_{\text{fixed}} \bigcup C_1, \\ 1 - \frac{1}{\pi}\cos^{-1}\Big(\frac{2R2^2_{\text{fixed}}}{R^2} - \frac{(R^2 - 2R2^2_{\text{fixed}} - v^2\tau_o^2)^2}{2v^2\tau_o^2 R^2} - 1\Big), & R - v\tau_o \le 2R2_{\text{fixed}} \le R \bigcup C_1, \\ 0, & 2R2_{\text{fixed}} < R - v\tau_o \bigcup C_1, \\ \frac{1}{\pi}\cos^{-1}\Big(1 - \frac{2v^2\tau_o^2}{R^2}\Big), & C_2, \end{cases} \tag{7.4.2}$$

where $R2_{\text{fixed}}$ is the distance between the MT location and the AP of the WLAN cell when a handover out of the WLAN occurs in the fixed RSS threshold based method. It is calculated by

$$R2_{\text{fixed}} = 10^{\frac{E_t - \text{RSS2}_{\text{fixed}}}{10\beta}}. \tag{7.4.3}$$

In the hysteresis based method [LLGD08], a handover out of the WLAN is triggered when the RSS from the WLAN is below a threshold plus a hysteresis, $\text{RSS2}_{\text{hyst}}+h_y$, where $\text{RSS2}_{\text{hyst}}$ is the RSS threshold and h_y is a constant representing the hysteresis.

Using Equation (6.3.12), the connection breakdown probability for the hysteresis based method is given by

$$P_{b_{\text{hyst}}} = \begin{cases} 1, & R < R2_{\text{hyst}} \bigcup C_1, \\ 1 - \frac{1}{\pi}\cos^{-1}\left(\frac{2R2_{\text{hyst}}^2}{R^2} - \frac{(R^2 - 2R2_{\text{hyst}}^2 - v^2\tau_o^2)^2}{2v^2\tau_o^2 R^2} - 1\right), & R - v\tau_o \leq 2R2_{\text{hyst}} \leq R \bigcup C_1, \\ 0, & 2R2_{\text{hyst}} < R - v\tau_o \bigcup C_1, \\ \frac{1}{\pi}\cos^{-1}\left(1 - \frac{2v^2\tau_o^2}{R^2}\right), & C_2, \end{cases} \tag{7.4.4}$$

The probabilities of connection breakdowns of the RSS threshold based ($R2_{\text{fixed}}$ = 130 m), hysteresis based ($R2_{\text{hyst}}$ = 120 m) and HTCE methods are shown in Figure 7.12. Since HNE is designed to keep the probability of connection breakdowns below preset levels, even though the velocity of the MT increases, the probabilities remain the same. As illustrated by the figure, for high velocities, HTCE yields much lower probability of connection breakdowns than the other two methods.

7.4.2 Simulation Results of HTCE

For simulation, three critical parametric quantities are examined:

1. **Connection breakdown percentage.** A connection breakdown occurs if a handover is not completed before the MT travels out of the coverage of the currently connected network (i.e., the time period between the handover triggering and WLAN exit points is less than the handover delay from the

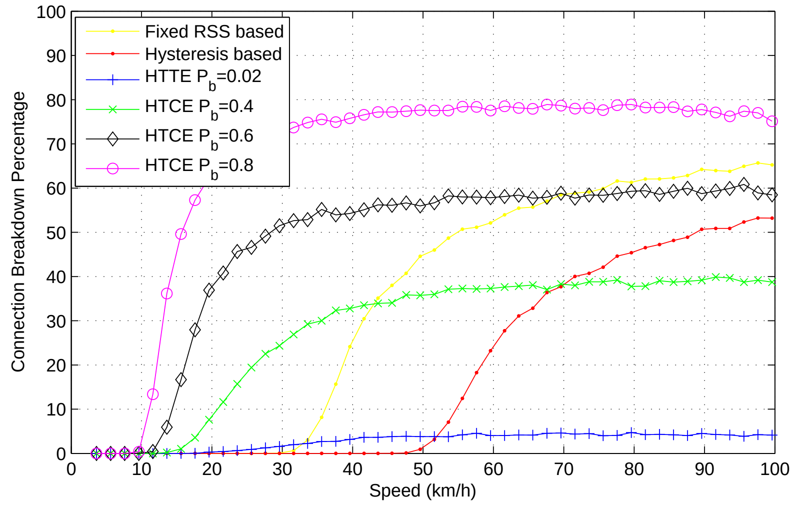

Figure 7.13: Connection breakdown percentage for the fixed RSS threshold based [VRWF03], hysteresis based [LLGD08] and HTCE methods.

WLAN to the cellular network). It is calculated as the percentage of the number of connection breakdowns in the total number of MT trajectories.

2. **WLAN usage.** For WLAN usage, two quantities are reported: total WLAN usage and breakdown-free WLAN usage. Total WLAN usage is the percentage of the time length that the MT is connected to the WLAN in the time length that the MT stays in the WLAN cell coverage. Breakdown-free WLAN usage is the same as total WLAN usage percentage, under the condition that the MT did not experience a connection breakdown in that trajectory. The value of the WLAN usage when a breakdown occurs is not included in the calculation of the breakdown-free WLAN usage.

3. **Handover triggering distance.** Handover triggering distance is the distance between the handover triggering point B and exit point P_o. This parameter is defined to enable comparison of the proposed algorithm against an ideal algorithm. An ideal algorithm would have knowledge of future channel behavior and trigger handover exactly at a point such that the handover is completed just as the MT leaves the coverage area. The handover triggering

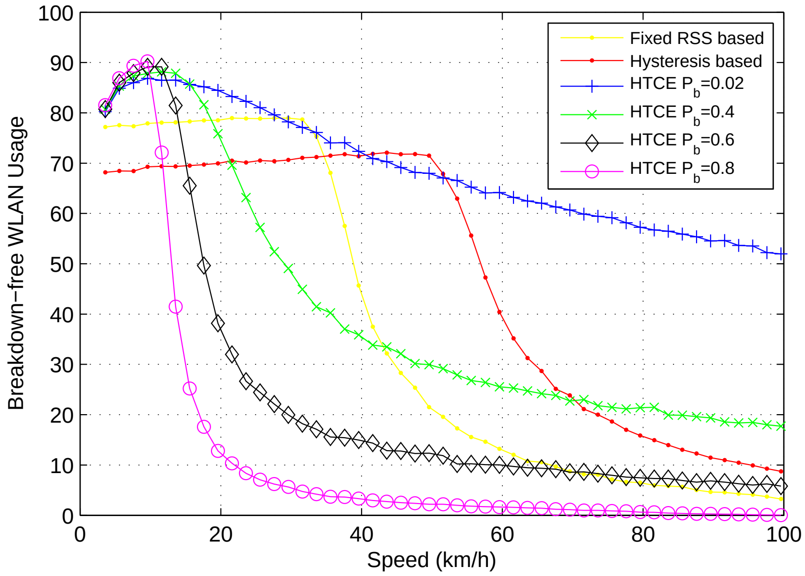

Figure 7.14: Breakdown-free WLAN usage for the fixed RSS threshold based [VRWF03], hysteresis based [LLGD08] and HTCE methods.

distance using the ideal algorithm would be the shortest possible distance leading to maximum breakdown-free WLAN usage.

MATLAB was used for the experiments, which generated 10000 random MT trajectories across the WLAN cell coverage area for speeds from 3.6 km/h to 100 km/h in 2 km/h increments. For each trajectory, a random WLAN cell entry point was chosen, and a uniformly distributed random angle between 0 and 2π was generated representing the movement direction of the MT.

Data for calculating the percentage of connection breakdowns were collected for RSS based method, hysteresis based method and HTCE with various breakdown tolerance settings (P_b), as shown in Figure 7.13. As can be seen from the figure, for HTCE, the breakdown percentage remains under the breakdown tolerances even for high speeds.

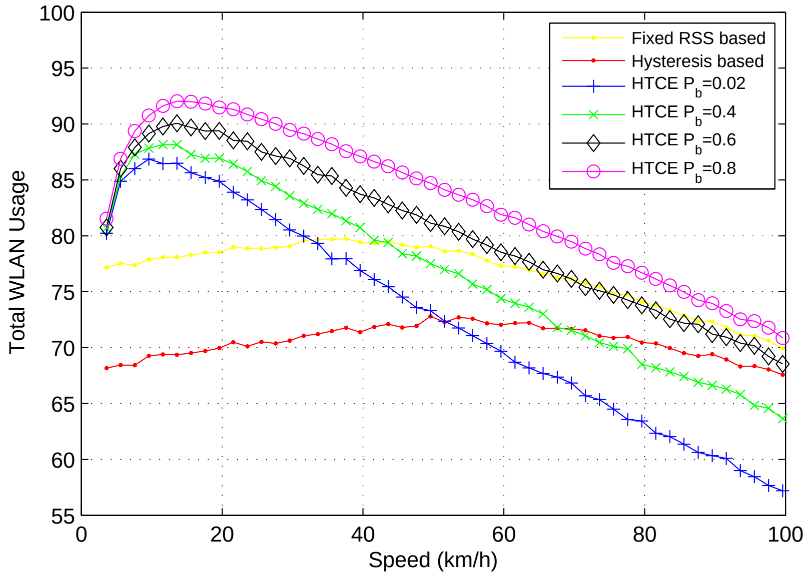

Figure 7.15: Total WLAN usage for the fixed RSS threshold based [VRWF03], hysteresis based [LLGD08] and HTCE methods.

The breakdown-free and total WLAN usage for RSS based method, hysteresis based method and HTCE with various breakdown tolerance settings are presented in Figures 7.14 and 7.15. Figure 7.14 shows that, for HTCE, breakdown-free WLAN usage remains above 50% when P_b is set to 0.02, which is much higher than the other two methods. As can be seen in Figure 7.15, for HTCE, total WLAN usage increases as higher values of P_b are applied. By adjusting P_b, an application running on the MT has the freedom on either maximizing WLAN usage or minimizing connection breakdown probability. For breakdown-sensitive applications such as voice, P_b could be set to a small value to maintain service continuity, while for other types of applications such as data, P_b could be set to a higher value to maximize the low-cost, high-bandwidth WLAN usage.

In Figure 7.16, handover triggering distances for ideal algorithm, RSS based method, hysteresis based method and HTCE with various breakdown tolerance settings are presented. As can be seen from the figure, the HTCE algorithm with the lowest breakdown tolerance ($P_b = 0.02$) has a very similar behavior, even though it

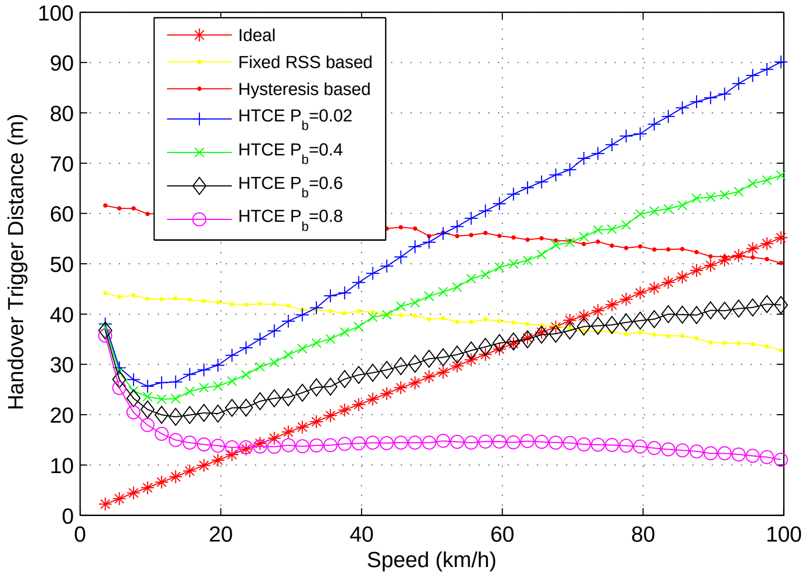

Figure 7.16: handover triggering distance for ideal algorithm, the fixed RSS threshold based [VRWF03], hysteresis based [LLGD08] and HTCE methods.

initiates the handover a bit earlier (approximately 20m) to accommodate for the potential estimation error based on RSS. With higher breakdown tolerance (for P_b value of 0.4, 0.6 and 0.8) HTCE takes higher risk of possible connection breakdown to increase the time spent in the WLAN coverage area. At P_b values of 0.6 and 0.8, HTCE delays handover beyond the ideal algorithm. These results demonstrate the flexibility of the proposed HTCE algorithm.

7.5 Conclusions

In this chapter, the performance evaluation of the proposed VHD algorithms has been presented. Through simulation experiments with MATLAB, it is demonstrated that the handover necessity estimation algorithm is able to reduce the number of handover failures and unnecessary handovers, the handover target selection algorithm is able to improve the user's satisfaction level, and the handover triggering condition estimation algorithm is able to offer the flexibility of choosing either for increased WLAN usage or reduced probability of connection breakdown. The combination of these three algorithms helps to optimize the VHD process.

Chapter 8

Concluding Remarks and Future Directions

8.1 Summary

The focus of the research project presented in this thesis report was to develop a vertical handover decision mechanism for 4G heterogeneous wireless networks. It was designed to be modular and can be implemented within the IEEE 802.21 MIH framework. The main aim of the scheme is to minimize communication interruptions due to handovers whilst maximizing the utilization of network resources in a cost effective way. In particular, the following requirements were taken into account:

- Minimize the probability of handover failures, unnecessary handovers and connection breakdowns in the event of a handover.
- Maximize the connection time with a preferred network.
- Maximize the user satisfaction level.

Three modules were proposed, namely, handover necessity estimation (HNE), handover target selection (HTS) and handover triggering condition estimation

Approach	Service continuity optimization	Network resource usage optimization	User satisfaction optimization	Adaptation to user mobility	Provision of user control over the tradeoff between service continuity and network resource usage
RSS based	Yes	No	No	Yes	No
Bandwidth based	No	Yes	No	No	No
Cost function based	No	Yes	Yes	No	No
Combination based	No	Yes	Yes	No	No
Proposed VHD scheme	Yes	Yes	Yes	Yes	Yes

Table 8.1: A comparative summary of VHD approaches.

(HTCE). HNE determines whether a handover is necessary when a MT enters the coverage of a WLAN cell. HTS selects a handover target when several candidate networks are available. HTCE determines the timing to initiate a handover from the currently connected WLAN, when the MT is leaving the WLAN cell's coverage area. These three modules work closely with each other. Depending on the movement of the MT, one of the modules is applied for helping the handover decision.

Table 8.1 compares the proposed VHD scheme with the surveyed VHD approaches presented in Chapter 2. As discussed in Section 2.5, RSS based approaches are targeted at optimizing the service continuity and can adapt to the user's mobility, whilst bandwidth based approaches are designed for optimizing the network resource usage. Cost function based and combination VHD approaches optimize both the network resource usage and user satisfaction, but fail to consider the service continuity or user movement. The VHD scheme proposed in this thesis is able to provide all the features mentioned above. It also offers users the flexibility of choosing between increased network resource usage and reduced service interruption, which are not provided in the other VHD methods.

8.2 Outcomes of the Research Project

This section recaps the major contributions of the work presented in this thesis.

Chapter 2 presented a critical review of related work in the area of VHD algorithms. Chapter 3 presented an overview of the proposed VHD scheme, connections between the three modules (HNE, HTS and HTCE) along with brief descriptions of each.

In Chapter 4, mathematical justification of the HNE module was presented. This method aims at minimizing handover failures and unnecessary handovers. HNE is composed of two algorithms, traveling time estimation and time threshold calculation. Traveling time estimation relies on RSS measurements and the speed of the MT. While the time threshold is calculated based on various network parameters such as tolerable handover failure probability or unnecessary handover probability, the radius of the WLAN cell and the handover latency.

In Chapter 5, mathematical justification of the HTS module was presented. This method is devised for maximizing the user satisfaction level by selecting the "best" network as the handover target among multiple candidate networks. HTS involves two algorithms, weights distribution and cost factor calculation. Weights of various network parameters are generated based on user preferences and the power level of the MT. Cost factors of candidate networks are calculated using a cost function. The network with the lowest cost factor is selected as the handover target.

In Chapter 6, mathematical justification of the HTCE module was presented. This method is devised to keep the connection breakdown probability below user specified limits under different MT speeds, as well as to provide the user with control over the tradeoff between connection breakdown probability and WLAN usage. HTCE calculates a RSS threshold for triggering a handover based on the estimated radius of the WLAN cell, handover latency, speed of the MT and connection breakdown tolerance.

An evaluation of the proposed VHD methods was presented in Chapter 7. MATLAB based simulation experiments demonstrated that:

- HNE reduced the number of handover failures and unnecessary handovers up to 80% and 70%,
- HTS increased the user's satisfaction level up to 50%, and
- HTCE was able to offer the flexibility of choosing between increased WLAN connection time and reduced probability of connection breakdown.

Effective vertical handover decision schemes remain as a crucial factor in effective realization of future 4G mobile technology relying on heterogeneous wireless access. This thesis presented a detailed outline of the major issues encountered in such an environment and conducted a systematic study for demonstration of feasible solutions towards fully integrated, efficient heterogeneous wireless networks. The combined VHD scheme is able to optimize the handover decision process for 4G mobile users.

8.3 Future Directions

In this thesis, required building blocks for optimization of the handover performance were developed, however, the measurement of application related performance parameters such as packet losses and throughput have yet to be implemented in the simulation. It is desirable to evaluate these parameters in order to thoroughly verify the proposed VHD scheme in practice. Ultimate evaluation of the proposed scheme should be conducted on dedicated testbeds in a laboratory environment.

Another future direction is to tune the proposed algorithms using real propagation measurement data, and thus the algorithms can be applied to various realistic conditions, e.g. propagation environments. This can potentially enhance the accuracy of the performance evaluation since WLANs are usually deployed in urban areas in which shadow fading is significant.

Further areas of research could include the effect of user velocities and propagation environments. Different mobility models and propagation models could be applied representing various user scenarios such as pedestrians in a campus, bicycle riders in an urban area and motorists in a suburban area.

Appendix A

Performance Evaluations in MATLAB

MATLAB programs written for the evaluation of the performance of proposed VHD methods are provided in this section. Programs for HNE, HTS and HTCE modules are listed separately, and flowcharts of each program are included.

A.1 Performance Evaluation of HNE

A.1.1 Theoretical Performance Evaluation of HNE

A flowchart of the MATLAB program for the theoretical performance evaluation of HNE is included in Figure A.1. The key portion of the MATLAB code is provided below.

```
%% This code is used to compare the proposed handover necessity estimation
%% (HNE) with other handover decision methods, including fixed RSS based
%% and hysteresis based methods.
clear all

%% Network parameters initialization
R=150;R_fixed=150;R_hyst=120;tau_i=2;tau_o=2;P_f=0.02;P_u=0.04;

%% Velocities of mobile terminals from 1 km/h to 100 km/h are included.
for i=1:100
    v0(i)=i;v(i)=v0(i).*5/18;
```

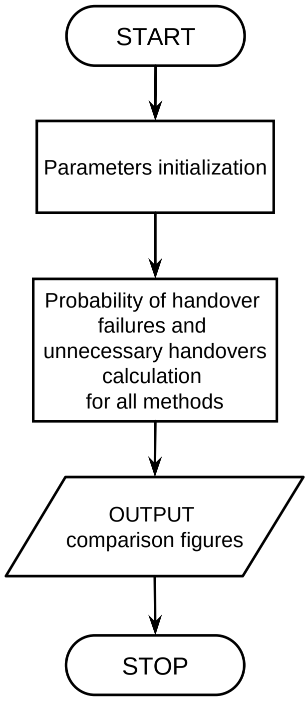

Figure A.1: Flowchart of the MATLAB program for the theoretical performance evaluation of HNE.

```
%% Probability of handover failures for the fixed RSS based method.
    if v(i)*tau_i>=2*R_fixed
        P_f_fixed(i)=1;
    else P_f_fixed(i)=2/pi*asin(v(i).*tau_i/2/R_fixed);
    end
%% Probability of unnecessary handovers for the fixed RSS based method.
    if v(i)*(tau_i+tau_o)>=2*R_fixed
        P_u_fixed(i)=1;
    else P_u_fixed(i)=2/pi*asin(v(i).*(tau_i+tau_o)/2/R_fixed);
    end
%% Hysteresis based method is similar to fixed RSS based method.

%% Probability of handover failures for HNE.
    P_f_HNE(i)=P_f;
%% Probability of unnecessary handovers for HNE.
    P_u_HNE(i)=P_u;
end
```

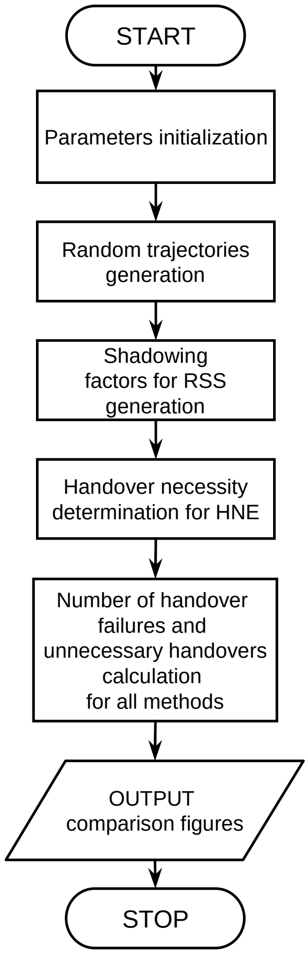

Figure A.2: Flowchart of the MATLAB program for the simulation based performance evaluation of HNE.

A.1.2 Simulation Based Performance Evaluation of HNE

A flowchart of the MATLAB program for the simulation based performance evaluation of HNE is included in Figure A.2. The key portion of the MATLAB code is provided below.

```
%% This code is used to compare the proposed handover necessity estimation
%% (HNE) with other handover decision methods, including fixed RSS based
%% and hysteresis based methods.
clear all

%% Network parameters initialization.
beta =3.5;P_tx=20;PL_ref=40;R=150;R_fixed=150;R_hyst=120;

```

```
%% Velocities of mobile terminals from 1 km/h to 100 km/h in 2 km/h
%% increments are simulated.
v0=[1:2:100];
for i=1:length(v0)
    v(i)=v0(i).*5/18;

  %% Generating 10000 trajectories with random entry and exist angles.
    for j = 1 : 10000
        theta_i(i,j)=2*pi*rand(1);
        theta_o(i,j)=2*pi*rand(1);
        theta(i,j)=theta_i(i,j)-theta_o(i,j);
        if theta(i,j) < 0
            theta(i,j) = abs(theta(i,j));
        end
        %% The length of sampling interval and number of RSS samples per
        %% second vary with the velocity of the mobile terminal.
        sample_interval(i)=0.001.*v(i);
        N(i)=floor(d_sample/sample_interval(i));
        sigma(i)=getsigma(sigma0,N(i));

        %% Theoretical traveling distance for the fixed RSS based method.
        d_fixed(i,j)=sqrt(2*R_fixed^2*(1-cos(theta(i,j))));
        l_OM_fixed(i,j) = sqrt(R_fixed^2-(d_fixed(i,j)/2).^2);
        %% Actual traveling distance for the fixed RSS based method.
        noise_fixed(i,j) = getnoise(sigma(i));
        RSS_fixed_est(i,j) = RSS_fixed+noise_fixed(i,j);
        R_fixed_est(i,j) = 10^((P_tx-PL_ref-RSS_fixed_est(i,j))/10/beta);
        d_fixed_est(i,j)=sqrt(2*R_fixed_est(i,j).^2.*(1-cos(theta(i,j))));
        l_OM_fixed_est(i,j) =
        sqrt(R_fixed_est(i,j).^2-(d_fixed_est(i,j)/2).^2);
        %% Calculate the number of handover failures and unnecessary
        %% handovers for the fixed RSS based method.
        if l_OM_fixed(i,j)<=R_fixed_est(i,j)
           totalnum_handover_fixed(i)=totalnum_handover_fixed(i)+1;
           if d_fixed_est(i,j)<v(i)*tau_i
               %% A handover failure (also an unnecessary handover) happens.
           num_failure_fixed(i)=num_failure_fixed(i)+1;
           num_unnecessary_fixed(i)=num_unnecessary_fixed(i)+1;
           elseif d_fixed_est(i,j)<v(i)*(tau_i+tau_o)
               %% An unnecessary handover happens.
           num_unnecessary_fixed(i)=num_unnecessary_fixed(i)+1;
           end
        end

        %% Hysteresis based method is similar to fixed RSS based method.

        %% Theoretical traveling distance for HNE
        d_HNE(i,j)=sqrt(2*R^2*(1-cos(theta(i,j))));
```

```
        l_OM_HNE(i,j) = sqrt(R^2-(d_HNE(i,j)/2).^2);
        %% Actual traveling distance for HNE
        noise_HNE(i,j) = getnoise(sigma(i));
        RSS_HNE_est(i,j) =RSS_HNE+noise_HNE(i,j);
        R_HNE_est(i,j) = 10^((P_tx-PL_ref-RSS_HNE_est(i,j))/10/beta);
        d_HNE_est(i,j)=sqrt(2*R_HNE_est(i,j).^2.*(1-cos(theta(i,j))));
        l_OM_HNE_est(i,j) = sqrt(R_HNE_est(i,j).^2-(d_HNE_est(i,j)/2).^2);
        %% Calculate the number of handover failures and unnecessary
        %% handovers for HNE
        if l_OM_HNE(i,j)<=R_HNE_est(i,j)
            %% If not, no handover happens.
           D1=2*R*sin(asin(v(i)*tau_i/2/R)-pi/2*P_f);
           %% Distance threshold for minimizing handover failures.
           D2=2*R*sin(asin(v(i)*(tau_i+tau_o)/2/R)-pi/2*P_u);
           %% Distance threshold for minimizing unnecessary handovers.
           if d_HNE(i,j)>=D1
              totalnum_handover_HNE_pf(i)=totalnum_handover_HNE_pf(i)+1;
              if d_HNE_est(i,j)<v(i)*tau_i
                 num_failure_HNE(i)=num_failure_HNE(i)+1;
                 %% Handover failure happens.
              end
           end
           if d_HNE(i,j)>=D2
              totalnum_handover_HNE_pu(i)=totalnum_handover_HNE_pu(i)+1;
              if d_HNE_est(i,j)<v(i)*(tau_i+tau_o)
                  %% Unnecessary handover happens.
                 num_unnecessary_HNE(i)=num_unnecessary_HNE(i)+1;
```

A.2 Simulation Based Performance Evaluation of HTS

A flowchart of the MATLAB program for the simulation based performance evaluation of HTS is included in Figure A.3. The key portion of the MATLAB code is provided below.

```
%% This code is used to compare the proposed handover target selection
%% (HTS) method with methods which consistently choose one access network,
%% under different user expected costs per minute. The other three
%% programs (which are for different user expected bandwidth, security
%% levels and power levels of the MT) are similar to this one and thus are
%% not included in the thesis.
clear all
```

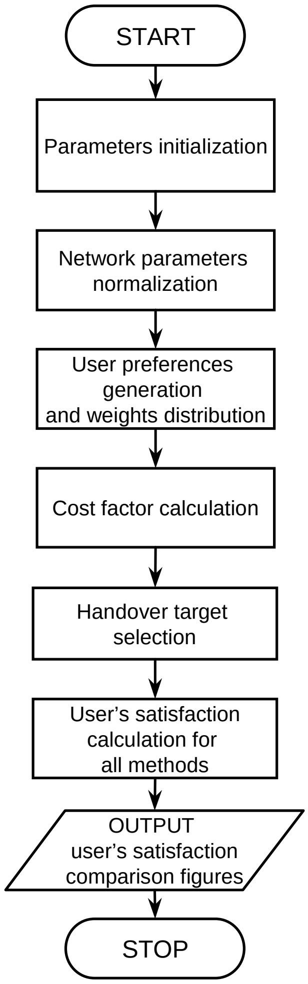

Figure A.3: Flowchart of the MATLAB program for the simulation based performance evaluation of HTS.

```
%% User expectations of monetary cost (variable), bandwidth (fixed) and
%% security (fixed).
exp_money=1:10;
exp_bandwidth=500;
exp_security=3;

%% Network parameters initialization.
bandwidth1=2000;money1=3;security1=1;power1=8;

%% Network parameters Normalization.
N_bandwidth_norm = (N_bandwidth_cost)/max(N_bandwidth_cost);

%% Weights distribution for HTS.
for i=1:length(exp_money)
for j=1:1000
i_bandwidth(i,j)=randint(1,1,4);
```

```
i_cost(i,j)=randint(1,1,4);
i_security(i,j)=randint(1,1,4);
if i_bandwidth(i,j)~=0 || i_cost(i,j)~=0 || i_security(i,j)~=0
     [ w_bandwidth(i,j),w_cost(i,j),w_security(i,j),w_power(i,j) ] =
     weights( i_bandwidth(i,j),i_cost(i,j),i_security(i,j),p_w );
else [ w_bandwidth(i,j),w_cost(i,j),w_security(i,j),w_power(i,j) ] =
    weights( 1,1,1,p_w );
end

%% Handover target selection based on the cost function for HTS.
cost_Net1(i,j)=N_bandwidth_cost(1)*w_bandwidth(i,j)+N_money(1)*w_cost(i,j)+
N_security_cost(1)*w_security(i,j)+N_power(1)*w_power(i,j);
%% Cost factors for network 2 and 3 are similar.
if min_cost(i,j)==cost_Net1(i,j)
    act_bandwidth(i,j)=N_bandwidth(1);
    act_money(i,j)=N_money(1);
    act_security(i,j)=N_security(1);
elseif min_cost(i,j)==cost_Net2(i,j)

%% Calculation of user's satisfaction.
satisfaction_HTS(i,j)=
(act_bandwidth(i,j)-exp_bandwidth)/act_bandwidth(i,j)*w_bandwidth(i,j) +
(exp_money(i)-act_money(i,j))/act_money(i,j)*w_cost(i,j) +
(act_security(i,j)-exp_security)/act_security(i,j)*w_security(i,j);
%% User's satisfaction for network 1, 2 and 3 is similar.
```

A.3 Performance Evaluation of HTCE

A.3.1 Theoretical Performance Evaluation of HTCE

A flowchart of the MATLAB program for the theoretical performance evaluation of HTCE is included in Figure A.4. The key portion of the MATLAB code is provided below.

```
%% This code is used to compare the proposed handover triggering condition
%% estimation (HTCE) with other handover triggering methods, including
%% fixed RSS based and hysteresis based methods.
clear all

%% Network parameters initialization.
R=150;R_fixed=130;R_hyst=120;tau_i=2;tau_o=2;P_b=0;

```

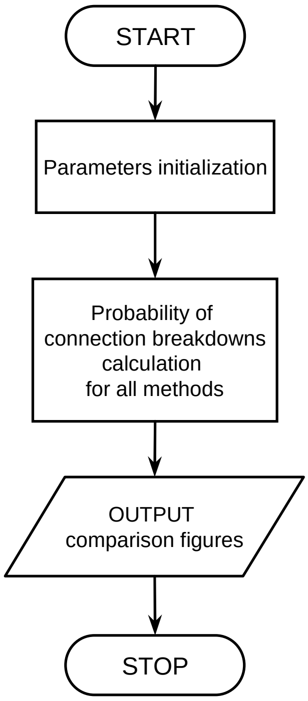

Figure A.4: Flowchart of the MATLAB program for the theoretical performance evaluation of HTCE.

```
%% Velocities of mobile terminals from 1 km/h to 100 km/h are included.
for i=1:100
    v0(i)=i;v(i)=v0(i).*5/18;

%% Probability of connection breakdowns for the fixed RSS based method.
    if v(i)*tau_o<=(R-R_fixed)
        P_b_fixed(i)=0;
    elseif v(i)*tau_o>=(sqrt(R^2-R_fixed^2))
        P_b_fixed(i)=1;
    else P_b_fixed(i)=1-1/pi*(acos(2*R_fixed^2/R^2-(R^2-R_fixed^2-(v(i)*
        tau_o).^2).^2/2./(v(i)*tau_o).^2/R^2-1));
    end

%% Hysteresis based method is similar to fixed RSS based method.

%% Probability of connection breakdowns for HTCE.
    P_b_HTCE(i)=P_b;
end
```

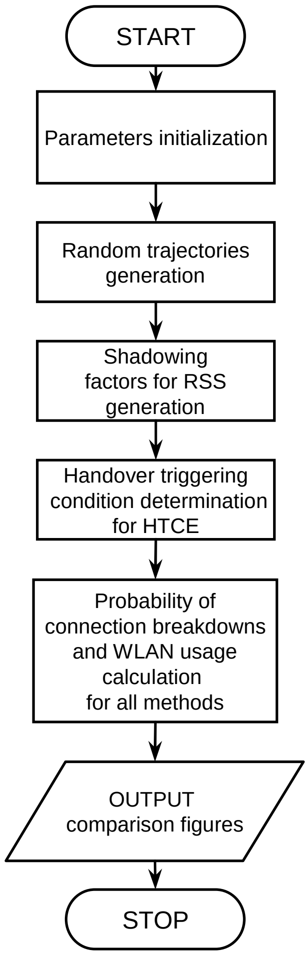

Figure A.5: Flowchart of the MATLAB program for the simulation based performance evaluation of HTCE.

A.3.2 Simulation Based Performance Evaluation of HTCE

A flowchart of the MATLAB program for the simulation based performance evaluation of HTCE is included in Figure A.5. The key portion of the MATLAB code is provided below.

```
%% This code is used to compare the proposed handover triggering condition
%% estimation (HTCE) with other handover triggering methods, including
%% fixed RSS based and hysteresis based methods.
clear all

%% Network parameters initialization.
R=150;sigma0=4.3;d_sample=30;P_tx=20;PL_ref=40;tau_i=2;tau_o=2;beta=3.5;

```

```
%% Four values of P_b are applied in the simulation: 0.02, 0.4, 0.6 and
%% 0.8.

%% Velocities of mobile terminals from 3.6 km/h to 100 km/h in 2 km/h
%% increments are simulated.
for ik = 1 :length(v)
    m=0;n=0;m1=0;m3=0;n1=0;n2=0;n3=0;m0=0;
    sample_interval(ik)=0.001.*v0(ik);
    N(ik)=floor(d_sample/sample_interval(ik));
    sigma(ik)=getsigma(sigma0,N(ik));
    errorratio(ik)=0.1./v(ik);

    %% Generating 10000 trajectories with random entry and exist angles.
    for ij = 1 : 10000
        breakdown_HTCE(i,ik,ij)=0;
        breakdown_GHO(ik,ij)=0;
        breakdown_fixed(ik,ij)=0;
        theta_i(ik,ij)=2*pi*rand(1);
        theta_o(ik,ij)=2*pi*rand(1);
        theta(ik,ij)=theta_i(ik,ij)-theta_o(ik,ij);
        if theta(ik,ij) < 0
            theta(ik,ij) = abs(theta(ik,ij));
        end
        L(ik,ij)=sqrt(2*R^2*(1-cos(theta(ik,ij))));
        Lh(ik,ij) = 0.5*L(ik,ij);
        L_OM(ik,ij) = sqrt(R^2-Lh(ik,ij)^2);

        %% Calculate the handover triggering distance of the ideal
        %% algorithm.
        r_ideal(ik,ij) = sqrt(L_OM(ik,ij)^2+(Lh(ik,ij)-v(ik)*tau_o)^2);
        noise_ideal(ik,ij) = getnoise(sigma(ik));
        RSS_ideal(ik,ij) = P_tx-PL_ref-10*beta*log10(r_ideal(ik,ij))+
        noise_ideal(ik,ij);
        r_ideal_est(ik,ij) = 10^((P_tx-PL_ref-RSS_ideal(ik,ij))/10/beta);
        m0=m0+1;
        if (theta(ik,ij)>=acos(1-v(ik)^2*(tau_i+tau_o)^2/2/R^2)) &
            (theta(ik,ij)<=2*pi-acos(1-v(ik)^2*(tau_i+tau_o)^2/2/R^2))
            if R < v(ik)*tau_o
                boundary_ideal(ik,ij) = R;
                wuse_ideal(ik) = wuse_ideal(ik)+L(ik,ij)/v(ik)-tau_i;
            else
        if r_ideal_est(ik,ij)>L_OM(ik,ij) & r_ideal_est(ik,ij)<R
        boundary_ideal(ik,ij) = Lh(ik,ij) - sqrt(r_ideal_est(ik,ij)^2-
        L_OM(ik,ij)^2);
        if boundary_ideal(ik,ij)>=tau_o*v(ik)
            wuse_eff_ideal(ik) = wuse_eff_ideal(ik)+L(ik,ij)/v(ik)-tau_i+
            tau_o- boundary_ideal(ik,ij)/v(ik);
            wuse_ideal(ik) = wuse_ideal(ik)+L(ik,ij)/v(ik)-tau_i+tau_o-
```

```
                boundary_ideal(ik,ij)/v(ik);
            end
        elseif r_ideal_est(ik,ij)>=R
            boundary_ideal(ik,ij) = 0;
            wuse_ideal(ik) = wuse_ideal(ik)+L(ik,ij)/v(ik)-tau_i;
        else
            boundary_ideal(ik,ij) = Lh(ik,ij);
            if (L(ik,ij)-tau_i*v(ik)) < v(ik)*tau_o
                wuse_ideal(ik) = wuse_ideal(ik)+L(ik,ij)/v(ik)-tau_i;
            else
                 wuse_ideal(ik) =wuse_ideal(ik)+tau_o;
                 wuse_eff_ideal(ik) =wuse_eff_ideal(ik)+tau_o;
            end
        end
        distance_ideal(ik) = distance_ideal(ik)+boundary_ideal(ik,ij);
            end
        end
        wlantime(ik) = wlantime(ik) + L(ik,ij)/v(ik);

        %% Calculate the probability of connection breakdowns, WLAN usage
        %% and handover triggering distance for HTCE.
        if (theta(ik,ij)>=acos(1-v(ik)^2*(tau_i+tau_o)^2/2/R^2)) &
            (theta(ik,ij)<=2*pi-acos(1-v(ik)^2*(tau_i+tau_o)^2/2/R^2))
            if R < v(ik)*tau_o
                num_break(i,ik,ij) = num_break(ik,ij)+1;
                wuse(i,ik) = wuse(i,ik)+L(ik,ij)/v(ik)-tau_i;
                L_bound_est(i,ik,ij) = R;
                L_bound_est_eff(i,ik,ij) = 0;
            else
                if v0(ik)<v_htce
                    r(ik) = R_hyst;
                else
                r(i,ik) = sqrt(v(ik)^2*tau_o^2+R^2-
                sqrt(2*(1-cos(pi-pi*Ph(i))))*v(ik)*tau_o*R);
                r(i,ik) = r(i,ik)-R*errorratio(ik);
                end
                L_bound(ik) = v(ik)*tau_o;
                l_OS1(ik,ij) = sqrt((v(ik).*0).^2+L_OM(ik,ij)^2);
                l_OS2(ik,ij) = sqrt((2*(v(ik)*0))^2+L_OM(ik,ij)^2);
                noise_S1(ik,ij) = getnoise(sigma(ik));
                noise_S2(ik,ij) = getnoise(sigma(ik));
                RSS_S1(ik,ij) = P_tx-PL_ref-10*beta*log10(l_OS1(ik,ij))+
                noise_S1(ik,ij);
                RSS_S2(ik,ij) = P_tx-PL_ref-10*beta*log10(l_OS2(ik,ij))+
                noise_S2(ik,ij);
              l_OS1_est(ik,ij) = 10^((P_tx-PL_ref-RSS_S1(ik,ij))/10/beta);
              l_OS2_est(ik,ij) = 10^((P_tx-PL_ref-RSS_S2(ik,ij))/10/beta);
                noise_r(ik,ij) = getnoise(sigma(ik));
```

```
                RSS_r(i,ik,ij) = P_tx-PL_ref-10*beta*log10(r(i,ik))+
                noise_r(ik,ij);
                r_est(i,ik,ij) = 10^((P_tx-PL_ref-RSS_r(i,ik,ij))/10/beta);
                d_R2(ik,ij)=sqrt(0.5*R^2*(1-cos(theta(ik,ij))));
                        if r_est(i,ik,ij)>=R
                            breakdown_HTCE(i,ik,ij)=1;
                           num_break(i,ik) = num_break(i,ik)+1;
                           wuse(i,ik) = wuse(i,ik)+L(ik,ij)/v(ik)-tau_i;
                           L_bound_est(i,ik,ij) = 0;
                           L_bound_est_eff(i,ik,ij) = 0;
                        elseif r_est(i,ik,ij)<L_OM(ik,ij)
                            if (L(ik,ij)-tau_i*v(ik)) < v(ik)*tau_o
                            breakdown_HTCE(i,ik,ij)=1;
                            num_break(i,ik) = num_break(i,ik)+1;
                            wuse(i,ik) = wuse(i,ik)+L(ik,ij)/v(ik)-tau_i;
                            L_bound_est(i,ik,ij) = L(ik,ij)-tau_i*v(ik);
                            L_bound_est_eff(i,ik,ij) = 0;
                            else
                                   wuse(i,ik) =wuse(i,ik)+tau_o;
                                   wuse_eff(i,ik) =wuse_eff(i,ik)+0+tau_o;
                                   L_bound_est(i,ik,ij) =
                                   L(ik,ij)-v(ik)*tau_i;
                                   L_bound_est_eff(i,ik,ij) =
                                   L(ik,ij)-v(ik)*tau_i;
                                   n1=n1+1;
                               end
                        else
                            d_r(i,ik,ij)=sqrt(r_est(i,ik,ij)^2-
                            L_OM(ik,ij)^2);
                            L_bound_est(i,ik,ij)=d_R2(ik,ij)-d_r(i,ik,ij);
                            if L_bound_est(i,ik,ij) ~= 0  &
                               L_bound_est(i,ik,ij) < v(ik)*tau_o
                               breakdown_HTCE(i,ik,ij)=1;
                               num_break(i,ik) = num_break(i,ik)+1;
                               wuse(i,ik) = wuse(i,ik)+L(ik,ij)/v(ik)-tau_i;
                               L_bound_est_eff(i,ik,ij) = 0;
                            elseif L_bound_est(i,ik,ij) ~= 0
                              wuse(i,ik) =wuse(i,ik)+L(ik,ij)/v(ik)-
                              tau_i+tau_o-L_bound_est(i,ik,ij)/v(ik);
                              wuse_eff(i,ik)=wuse_eff(i,ik)+L(ik,ij)/v(ik)-
                              tau_i+tau_o-L_bound_est(i,ik,ij)/v(ik);
                              L_bound_est_eff(i,ik,ij) =
                              L_bound_est(i,ik,ij);
                              n1=n1+1;
                            end
                        end
            end
            esterror(i,ik,ij) = abs(L_bound_est(i,ik,ij)-
```

```
            L_bound(ik))/L_bound(ik);
                        esterror_sum(i,ik) = esterror_sum(i,ik)+
                        esterror(i,ik,ij);
                        m = m+1;
                        distance_HTCE(i,ik) = distance_HTCE(i,ik)+
                        L_bound_est(i,ik,ij);
                        distance_HTCE_eff(i,ik) = distance_HTCE_eff(i,ik)+
                        L_bound_est_eff(i,ik,ij);
        end

        %% Calculate the probability of connection breakdowns and WLAN
        %% usage for the hysteresis based method.
            if (theta(ik,ij)>=acos(1-v(ik)^2*(tau_i+tau_o)^2/2/R^2)) &
                (theta(ik,ij)<=2*pi-acos(1-v(ik)^2*(tau_i+tau_o)^2/2/R^2))
            noise_rHY(ik,ij) = getnoise(sigma(ik));
            RSS_rHY(ik,ij) = P_tx-PL_ref-10*beta*log10(R_hyst)+
            noise_rHY(ik,ij);
            rHY_est(ik,ij) = 10^((P_tx-PL_ref-RSS_rHY(ik,ij))/10/beta);
            Min_err_sum(ik) = 0;
            Min_err(ik,ij) = abs(rHY_est(ik,ij)-R_hyst)/R_hyst;
            Min_err_sum(ik) = Min_err_sum(ik)+Min_err(ik,ij);
            m1 = m1+1;
            d_R2_Min(ik,ij)=sqrt(0.5*R^2*(1-cos(theta(ik,ij))));
            if rHY_est(ik,ij)>=R
                breakdown_GHO(ik,ij)=1;
                num_break_Min(ik) = num_break_Min(ik)+1;
                wuse_Min(ik) = wuse_Min(ik)+L(ik,ij)/v(ik)-tau_i;
                L_bound_est_GHO(ik,ij) = 0;
                L_bound_est_GHO_eff(ik,ij) = 0;
            elseif rHY_est(ik,ij)<=L_OM(ik,ij) & rHY_est(ik,ij) ~=0
                   if (L(ik,ij)-tau_i*v(ik)) < v(ik)*tau_o
                   breakdown_GHO(ik,ij)=1;
                   num_break_Min(ik) = num_break_Min(ik)+1;
                   wuse_Min(ik) = wuse_Min(ik)+L(ik,ij)/v(ik)-tau_i;
                   L_bound_est_GHO(ik,ij) = L(ik,ij)-tau_i*v(ik);
                   L_bound_est_GHO_eff(ik,ij) = 0;
                   else wuse_Min(ik) = wuse_Min(ik)+0+tau_o;
                       wuse_eff_Min(ik) = wuse_eff_Min(ik)+0+tau_o;
                       L_bound_est_GHO(ik,ij) = L(ik,ij)-tau_i*v(ik);
                       L_bound_est_GHO_eff(ik,ij) = L_bound_est_GHO(ik,ij);
                       n2=n2+1;
                   end
            else
            d_r_Min(ik,ij)=sqrt(rHY_est(ik,ij)^2-R^2+d_R2_Min(ik,ij)^2);
            d2_Min(ik,ij)=d_R2_Min(ik,ij)-d_r_Min(ik,ij);
            L_bound_est_GHO(ik,ij) = d2_Min(ik,ij);
            if d2_Min(ik,ij) ~= 0  &  d2_Min(ik,ij) < v(ik)*tau_o
                breakdown_GHO(ik,ij)=1;
```

```
                    num_break_Min(ik) = num_break_Min(ik)+1;
                    wuse_Min(ik) = wuse_Min(ik)+L(ik,ij)/v(ik)-tau_i;
                    L_bound_est_GHO_eff(ik,ij) = 0;
                elseif d2_Min(ik,ij) ~= 0
                    wuse_Min(ik) =wuse_Min(ik)+L(ik,ij)/v(ik)-tau_i+tau_o-
                    d2_Min(ik,ij)/v(ik);
                    wuse_eff_Min(ik) =wuse_eff_Min(ik)+L(ik,ij)/v(ik)-
                    tau_i+tau_o-d2_Min(ik,ij)/v(ik);
                    L_bound_est_GHO_eff(ik,ij) = L_bound_est_GHO(ik,ij);
                    n2=n2+1;
                end
                end
                distance_GHO(ik) = distance_GHO(ik)+L_bound_est_GHO(ik,ij);
                distance_GHO_eff(ik) = distance_GHO_eff(ik)+
                L_bound_est_GHO_eff(ik,ij);
                end

            %% Fixed RSS based method is similar to hysteresis based method.
```

Bibliography

[Ada01] F. Adachi. Wireless past and future – evolving mobile communications systems. *IEICE Trans. Fundamentals*, E84-A(1):55–60, 2001.

[AH05] W. Ajib and D. Haccoun. An overview of scheduling algorithms in MIMO-based fourth-generation wireless systems. *IEEE Network*, 19(5):43–48, September 2005.

[AMH+99] I. F. Akyildiz, J. McNair, J. S. M. Ho, H. Uzunalioglu, and W. Wang. Mobility management in next-generation wireless systems. *Proceedings of the IEEE*, 87(8):1374–1384, August 1999.

[AMMC02] J. Al-Muhtadi, D. Mickunas, and R. Campbell. A lightweight reconfigurable security mechanism for 3G/4G mobile devices. *IEEE Wireless Communications*, 9(2):60–65, August 2002.

[AW02] I. F. Akyildiz and W. Wang. A dynamic location management scheme for next-generation multitier PCS systems. *IEEE Transactions on Wireless Communications*, 1(1):178–189, Jan 2002.

[AXM04] I. F. Akyildiz, J. Xie, and S. Mohanty. A survey of mobility management in next-generation all-IP-based wireless systems. *IEEE Wireless Communications*, 11(4):16–28, 2004.

[Bar04] F. Barcelo. Performance analysis of handoff resource allocation strategies through the state-dependent rejection scheme. *IEEE Transactions on Wireless Communications*, 3(3):900–909, 2004.

[Bea84] G. W. Beakley. Overview of commercial satellite communications. *IEEE Transactions on Aerospace and Electronic Systems*, AES-20(4):455–464, 1984.

[BHPC04] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa. Stochastic properties of the random waypoint mobility model. *Wireless Networks*, 10(5):555–567, 2004.

[BI04] S. Balasubramaniam and J. Indulska. Vertical handover supporting pervasive computing in future wireless networks. *Computer Communications*, 27(8):708–719, May 2004.

[BK10] A. Benmimoune and M. Kadoch. Vertical handoff between UMTS and WLAN. In *Proceedings of the 4th international conference on Communications and information technology (CIT'10)*, pages 131–140, Corfu Island, Greece, 2010.

[BRA01] IST-1999-10050 BRAIN. BRAIN architecture specifications and models, BRAIN functionality and protocol specification, March 2001.

[Cal97] M. H. Callendar. International mobile telecommunications-2000 standards efforts of the ITU. *IEEE Personal Communications*, 4(4):6–7, August 1997.

[CC08] B.-J. Chang and J.-F. Chen. Cross-layer-based adaptive vertical handoff with predictive RSS in heterogeneous wireless networks. *IEEE Transactions on Vehicular Technology*, 57(6):3679–3692, 2008.

[CCHL07] C. Chi, X. Cai, R. Hao, and F. Liu. Modeling and analysis of handover algorithms. In *Proceedings of the 2007 IEEE Global Telecommunications Conference (GLOBECOM'07)*, pages 4473–4477, Washington, DC, USA, November 2007.

[CGS94] G.E. Corazza, D. Giancristofaro, and F. Santucci. Characterization of handover initialization in cellular mobile radio networks. In *Proceedings of the 44th Vehicular Technology Conference (VTC'94)*, pages 1869–1872, Stockholm, Sweden, June 1994.

[CLB07] Y.-J. Choi, K. B. Lee, and S. Bahk. All-IP 4G network architecture for efficient mobility and resource management. *IEEE Wireless Communications*, 14(2):42–46, April 2007.

[CLH04] W.-T. Chen, J.-C. Liu, and H.-K. Huang. An adaptive scheme for vertical handoff in wireless overlay networks. In *Proceedings of the Tenth International Conference on Parallel and Distributed Systems (ICPADS'04)*, Newport Beach, California, USA, July 2004.

[CRMRS99] F. Cortes-Rodriguez, D. Munoz-Rodriguez, and R. Soto. Position location assisted multi-valued logic handoff algorithm. In *Proceedings of the 50th Vehicular Technology Conference (VTC'99 - Fall)*, pages 775–779, Amsterdam, the Netherlands, September 1999.

[CS05] W.-T. Chen and Y.-Y. Shu. Active application oriented vertical handoff in next-generation wireless networks. In *Proceedings of the 2005 IEEE Wireless Communications and Networking Conference (WCNC'05)*, pages 1383–1388, New Orleans, Louisiana, March 2005.

[CSC+04] L.-J. Chen, T. Sun, B. Chen, V. Rajendran, and M. Gerla. A smart decision model for vertical handoff. In *Proceedings of the 4th International Workshop on Wireless Internet and Reconfigurability (ANWIRE'04)*, pages 653–658, Athens, Greece, May 2004.

[CSH+01] P. M. L. Chan, R. E. Sheriff, Y. F. Hu, P. Conforto, and C. Tocci. Mobility management incorporating fuzzy logic for a heterogeneous IP environment. *IEEE Communications Magazine*, 39(12):42–51, 2001.

[CYA10] C. Ceken, S. Yarkan, and H. Arslan. Interference aware vertical handoff decision algorithm for quality of service support in wireless heterogeneous networks. *Computer Networks*, 54(5):726 – 740, 2010.

[DB98] A. Davies and T. Brady. Policies for a complex product system. *Futures*, 30(4):293 – 304, 1998.

[DDF+07] A. Dutta, S. Das, D. Famolari, Y. Ohba, K. Taniuchi, V. Fajardo, R. M. Lopez, T. Kodama, and H. Schulzrinne. Seamless proactive handover across heterogeneous access networks. *Wireless Personal Communications*, 43(3):837–855, 2007.

[DGIM02] M. Datar, A. Gionis, P. Indyk, and R. Motwani. Maintaining stream statistics over sliding windows (extended abstract). In *SODA '02: Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 635–644, San Francisco, California, January 2002.

[DLOBS+08] A. De La Oliva, A. Banchs, I. Soto, T. Melia, and A. Vidal. An overview of IEEE 802.21: media-independent handover services. *IEEE Wireless Communications*, 15(4):96–103, August 2008.

[DPSB08] E. Dahlman, S. Parkvall, J. Skold, and P. Beming. *3G Evolution, Second Edition: HSPA and LTE for Mobile Broadband*. Academic Press, 2008.

[DStP03] IEEE802.11 WG Draft Supplement to Part 11. wireless medium access control (MAC) and physical layer (PHY) specification: Medium access control (MAC) enhancements for Quality of Service (QoS). *IEEE Standard 802.11e/D4.3*, May 2003.

[DSVK07] S. Dekleva, J. P. Shim, U. Varshney, and G. Knoerzer. Evolution and emerging issues in mobile wireless networks. *Communications of the ACM*, 50(6):38–43, 2007.

[DY05] D.-J. Deng and H.-C. Yen. Quality-of-Service provisioning system for multimedia transmission in IEEE 802.11 wireless LANs. *IEEE Journal on Selected Areas in Communications*, 23(6):1240–1252, 2005.

[Eri09] Ericsson. LTE - an introduction, 2009.

[FFF+06] S. Frattasi, H. Fathi, F. H. P. Fitzek, R. Prasad, and M. D. Katz. Defining 4g technology from the users perspective. *IEEE Network*, 20(1):35–41, January 2006.

[GAM05] V. Gazis, N. Alonistioti, and L. Merakos. Toward a generic "always best connected" capability in integrated WLAN/UMTS cellular mobile networks (and beyond). *IEEE Wireless Communications*, 12(3):20–29, June 2005.

[Gar02] L. Garber. Will 3G really be the next big wireless technology? *Computer*, 35(1):26–32, January 2002.

[GB06] M. Ghaderi and R. Boutaba. Call admission control in mobile cellular networks: a comprehensive survey: Research articles. *Wireless Communications and Mobile Computing*, 6(1):69–93, February 2006.

[GG05] F. Gustafsson and F. Gunnarsson. Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements. *IEEE Signal Processing Magazine*, 22(4):41–53, July 2005.

[GGZZ04] C. Guo, Z. Guo, Q. Zhang, and W. Zhu. A seamless and proactive end-to-end mobility solution for roaming across heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 22(5):834–848, 2004.

[GKCA08] A. George, A. Kumar, D. Cavalcanti, and D. P. Agrawal. Protocols for mobility management in heterogeneous multi-hop wireless networks. *Pervasive and Mobile Computing*, 4(1):92–116, 2008.

[Gol05] A. Goldsmith. *Wireless Communications*. Cambridge University Press, New York, NY, 2005.

[GS02] F. Graziosi and F. Santucci. A general correlation model for shadow fading in mobile radio systems. *IEEE Communications Letters*, 6(3):102–104, Mar 2002.

[Gud91] M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *IEEE Electronics Letters*, 27(23):2145 –2146, November 1991.

[GV03] F. Ghys and A. Vaaraniemi. Component-based charging in a next-generation multimedia network. *IEEE Communications Magazine*, 41(1):99–102, January 2003.

[GZX05] Q. Guo, J. Zhu, and X. Xu. An adaptive multi-criteria vertical handoff decision algorithm for radio heterogeneous network. In *Proceedings of the 2005 IEEE International Conference on Communications (ICC'05)*, pages 2769–2773, Seoul, Korea, May 2005.

[H+02] M.-J. Ho et al. RF challenges for 2.4 and 5 GHz WLAN deployment and design. In *Proceedings of the 2002 IEEE Wireless Communications and Networking Conference (WCNC'02)*, pages 783–788, Orlando, FL, USA, March 2002.

[HAK+09] D. Han, D. Andersen, M. Kaminsky, K. Papagiannaki, and S. Seshan. Access point localization using local signal strength gradient. In *Passive and Active Network Measurement*, volume 5448, chapter 10, pages 99–108. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[HHK06] S. Hussain, Z. Hamid, and N. S. Khattak. Mobility management challenges and issues in 4G heterogeneous networks. In *Proceedings of the first international conference on Integrated internet ad hoc and sensor networks (InterSense'06)*, Nice, France, May 2006. ACM.

[Hig08] S. Higginbotham. LTE vs WiMAX: A little 4G sibling rivalry, http://www.goinglte.com/lte-vs-wimax-a-little-4g-sibling-rivalry/, 2008.

[HNH06] A. Hasswa, N. Nasser, and H. Hassanein. Tramcar: A context-aware cross-layer architecture for next generation heterogeneous wireless networks. In *Proceedings of the 2006 IEEE International Conference on Communications (ICC'06)*, pages 240–245, Istanbul, Turkey, June 2006.

[HO06] J. Hou and D. C. O'Brien. Vertical handover-decision-making algorithm using fuzzy logic for the integrated radio-and-OW system. *IEEE Transactions on Wireless Communications*, 5(1):176–185, January 2006.

[Huo99] H. Huomo. Fourth generation mobile. In *Proceedings of the 4^{th} ACTS Mobile Communications Summit 1999 (AMOS'99)* [zz:99].

[HY03] S. Y. Hui and K. H. Yeung. Challenges in the migration to 4G mobile systems. *IEEE Communications Magazine*, 41(12):54–59, 2003.

[IEE01] IEEE Communications Society. *Proceedings of the IEEE Global Telecommunications Conference GLOBECOM'01*, San Antonio, TX, USA, 2001.

[KJC$^+$03] Y. Kim, B. J. Jeong, J. Chung, C.-S. Hwang, J. S. Ryu, Ki-Ho Kim, and Y. K. Kim. Beyond 3G: vision, requirements, and enabling technologies. *IEEE Communications Magazine*, 41(3):120–124, March 2003.

[KK04] M. Koutsopoulou and A. Kaloxylos. A holistic solution for charging, billing and accounting in 4G mobile systems. In *Proceedings of the 59^{rd} Vehicular Technology Conference (VTC'04 - Spring)*, Los Angeles, CA, USA, September 2004.

[KKP08] M. Kassar, B. Kervella, and G. Pujolle. An overview of vertical handover decision strategies in heterogeneous wireless networks. *Computer Communications*, 31(10):2607–2620, 2008.

[KLR02] D. N. Knisely, Q. Li, and N. S. Ramesh. CDMA2000: A third-generation radio transmission technology. *Bell Labs Technical Journal*, 3(3):63–78, August 2002.

[KUKR05] S. Kouhbor, J. Ugon, A. Kruger, and A. Rubinov. Optimal placement of access point in WLAN based on a new algorithm. In *ICMB '05: Proceedings of the International Conference on Mobile Business*, pages 592–598, Sydney Australia, 2005. IEEE.

[LCCS05] C. W. Lee, Li M. Chen, M. C. Chen, and Y. S. Sun. A framework of handoffs in wireless overlay networks based on mobile IPv6. *IEEE Journal on Selected Areas in Communications*, 23(11):2118–2128, 2005.

[LHH06] D. Lee, Y. Han, and J. Hwang. QoS-based vertical handoff decision algorithm in heterogeneous systems. In *Proceedings of the IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)*, pages 1–5, Helsinki, Finland, September 2006.

[LLGD08] M. Liu, Z. Li, X. Guo, and E. Dutkiewicz. Performance analysis and optimization of handoff algorithms in heterogeneous wireless networks. *IEEE Transactions on Mobile Computing*, 7(7):846–857, 2008.

[LLGL06] M. Liu, Z.-C. Li, X.-B. Guo, and H.-Y. Lach. Design and evaluation of vertical handoff decision algorithm in heterogeneous wireless networks. In *Proceedings of the 2006 IEEE International Conference on Networks (ICON'06)*, pages 1–6, Singapore, September 2006.

[LSG01] S.-J. Lee, W. Su, and M. Gerla. Wireless ad hoc multicast routing with mobility prediction. *Mobile Networks and Applications*, 6(4):351–360, 2001.

[LSK+09] S. Lee, K. Sriram, K. Kim, Y. H. Kim, and N. Golmie. Vertical handoff decision algorithms for providing optimized performance in heterogeneous wireless networks. *IEEE Transactions on Vehicular Technology*, 58(2):865–881, 2009.

[LTD06] H. Liao, L. Tie, and Z. Du. A vertical handover decision algorithm based on fuzzy control theory. In *Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*, pages 309–313, Hangzhou, China, June 2006.

[LWL08] Tsungnan Lin, Chiapin Wang, and Po-Chiang Lin. A neural-network-based context-aware handoff algorithm for multimedia computing. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 4(3):1–23, 2008.

[Lyl09] A. Lyle. Clear, first 4G network launched, http://www.neowin.net/news/main/09/01/10/clear-first-4g-network-launched-2, 2009.

[MA06] S. Mohanty and I. F. Akyildiz. A cross-layer (layer $2 + 3$) handover management protocol for next-generation wireless systems. *IEEE Transactions on Mobile Computing*, 5:1347–1360, October 2006.

[MAT92] MATLAB. *User's Guide*. The MathWorks, Inc., Natick, MA, 1992.

[Mis04] A. R. Mishra. *Fundamentals of cellular network planning and optimisation: 2G/2.5G/3G... Evolution to 4G*. Wiley, Chichester, West Sussex, England, 2004.

[MK00] W. Mohr and W. Konhauser. Access network evolution beyond third generation mobile communications. *IEEE Wireless Communications Magazine*, 38(12):122–133, 2000.

[Moh06] S. Mohanty. A new architecture for 3G and WLAN integration and inter-system handover management. *Wireless Networks*, 12(6):733–745, 2006.

[MOYU05] T. Miki, T. Ohya, H. Yoshino, and N. Umeda. The overview of the 4th generation mobile communication system. In *Proceedings of the Fifth International Conference on Information, Communications and Signal Processing (ICICS 2005)*, Bangkok, Thailand, December 2005.

[MP05] J.-P. Makela and K. Pahlavan. Performance of neural network handoff algorithm under varying mobile velocities. In *Proceedings of the 2005 Finnish Signal Processing Symposium - FINSIG'05*, pages 42–45, Kuopio, Finland, August 2005.

[MRD08] A. Misra, A. Roy, and S. K. Das. Information-theory based optimal location management schemes for integrated multi-system wireless networks. *IEEE/ACM Transactions on Networking*, 16(3):525–538, 2008.

[Mur99] K. Murota. Mobile communications trends in Japan and DoCoMo's activities towards 21st century. In *Proceedings of the 4^{th} ACTS Mobile Communications Summit 1999 (AMOS'99)* [zz:99].

[MZ04] J. McNair and F. Zhu. Vertical handoffs in fourth-generation multinetwork environments. *IEEE Wireless Communications*, 11(3):8–15, June 2004.

[NGAM07] N. Nasser, S. Guizani, and E. Al-Masri. Middleware vertical handoff manager: A neural network-based solution. In *Proceedings of the 2007 IEEE International Conference on Communications (ICC'07)*, pages 5671–5676, Glasgow, Scotland, June 2007.

[NH05] D. Niyato and E. Hossain. Call admission control for QoS provisioning in 4G wireless networks: issues and approaches. *IEEE Network*, 19(5):5–11, September 2005.

[NHH06a] N. Nasser, A. Hasswa, and H. Hassanein. Handoffs in fourth generation heterogeneous networks. *IEEE Communications Magazine*, 44:96–103, October 2006.

[NHH06b] N. Nasser, A. Hasswa, and H. Hassanein. Handoffs in fourth generation heterogeneous networks. *IEEE Communications Magazine*, 44(10):96–103, 2006.

[NVAGD07] Q.-T. Nguyen-Vuong, N. Agoulmine, and Y. Ghamri-Doudane. Terminal-controlled mobility management in heterogeneous wireless networks. *IEEE Communications Magazine*, 45(4):122–129, 2007.

[NWDZ05] J. Nie, J. Wen, Q. Dong, and Z. Zhou. A seamless handoff in IEEE 802.16a and IEEE 802.11n hybrid networks. In *Proceedings of the 2005 International Conference on Communications, Circuits and Systems (ICCCAS'05)*, pages 383–387, Hong Kong, China, May 2005.

[Pap65] A. Papoulis. *Probability, Random Variables, and Stochastic Processes. 1st ed.* McGraw-Hill, New York, 1965.

[PD02] G. Patel and S. Dennett. The 3GPP and 3GPP2 movements toward an all-IP mobile network. *IEEE Personal Communications*, 7(4):62–64, August 2002.

[Per00] J.M. Pereira. Fourth generation: now, it is personal! In *Proceedings of the IEEE 11th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'2000)*, London, UK, September 2000.

[PKH$^+$00] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. P. Makela, R. Pichna, and J. Vallstron. Handoff in hybrid mobile data networks. *IEEE Personal Communications*, 7(2):34–47, 2000.

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

[PP07] Y. Park and T. Park. A survey of security threats on 4G networks. In *Proceedings of the 2007 IEEE Global Telecommunications Conference (GLOBECOM'07)*, Washington, DC, USA, November 2007. IEEE Communications Society.

[PYK$^+$03] H. S. Park, H. S. Yoon, T. H. Kim, J. S. Park, M. S. Duo, and J. Y. Lee. Vertical handoff procedure and algorithm between IEEE802.11 WLAN and CDMA cellular network. *Mobile Communications*, pages 103–112, 2003.

[Rap02] Theodore S. Rappaport. *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall PTR, 2 edition, January 2002.

[Ric00] K. Richardson. UMTS overview. *Electronics and Communication Engineering Journal*, 12(3):93–100, 2000.

[SBB+06] J. Schuz, E. Bohler, G. Berg, B. Schlehofer, I. Hettinger, K. Schlaefer, J. Wahrendorf, K. Kunna-Grass, and M. Blettner. Cellular phones, cordless phones, and the risks of glioma and meningioma (interphone study group, germany). *American Journal of Epidemiology Advance Access*, 163(6):512–520, January 2006.

[SGR02] Sairam, N. Gunasekaran, and S. R. Redd. Bluetooth in wireless communication. *IEEE Communications Magazine*, 40(6):90–96, 2002.

[Sin07] B. Singh. An improved handover algorithm based on signal strength plus distance for interoperability in mobile cellular networks. *Wireless Personal Communications*, 43(3):879–887, November 2007.

[SNLW08] E. Stevens-Navarro, Y. Lin, and V. W. S. Wong. An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks. *IEEE Transactions on Vehicular Technology*, 57(2):1243–1254, 2008.

[SNW06] E. Stevens-Navarro and V. W. S. Wong. Comparison between vertical handoff decision algorithms for heterogeneous wireless networks. In *Proceedings of the 63rd Vehicular Technology Conference (VTC'06 - Spring)*, pages 947–951, Melbourne, Australia, May 2006.

[SOMGCV+03] E. Soria-Olivas, J.D. Martin-Guerrero, G. Camps-Valls, A.J. Serrano-Lopez, J. Calpe-Maravilla, and L. Gomez-Chova. A low-complexity fuzzy activation function for artificial neural networks. *IEEE Transactions on Neural Networks*, 14(6):1576–1579, November 2003.

[SSH01] J.-Z. Sun, J. Sauvola, and D. Howie. Features in future: 4G visions from a technical perspective. In *Proceedings of the IEEE Global Telecommunications Conference GLOBECOM'01* [IEE01].

[STT+02] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj. A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results. *IEEE Communications Magazine*, 40(9):143–149, 2002.

[Stü01] G. L. Stüber. *Principles of Mobile Communication*. Kluwer Academic Publishers, Norwell, MA, 2001.

[SZ06] F. Siddiqui and S. Zeadally. Mobility management across hybrid wireless networks: Trends and challenges. *Computer Communications*, 29(9):1363–1385, May 2006.

[Tac02] K. Tacikawa. *W-CDMA: Mobile Communications System*. Wiley, New York, NY, 2002.

[Tel99] E. Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, 10(6):585–595, 1999.

[TOF+09] K. Taniuchi, Y. Ohba, V. Fajardo, S. Das, M. Tauil, Y.-H. Cheng, A. Dutta, D. Baker, M. Yajnik, and D. Famolari. IEEE 802.21: Media independent handover: Features, applicability, and realization. *IEEE Communications Magazine*, 47(1):112–120, 2009.

[TPS08] R. Tawil, G. Pujolle, and O. Salazar. A vertical handoff decision scheme in heterogeneous wireless systems. In *Proceedings of the*

67^{th} Vehicular Technology Conference (VTC'08 - Spring), pages 2626–2630, Marina Bay, Singapore, April 2008.

[v.503] 3GPP TS 23.107 v.5.9.0. Quality of service (QoS) concept and architecture, June 2003.

[VJ01] U. Varshney and R. Jain. Issues in emerging 4G wireless networks. *Computer*, 34(6):94–96, 2001.

[VRWF03] V. K. Varma, S. Ramesh, K. D. Wong, and J. A. Friedhoffer. Mobility management in integrated UMTS/WLAN networks. In *Proceedings of the 2003 IEEE International Conference on Communications ICC'03*, pages 1048–1053, Anchorage, Alaska, USA, 2003. IEEE Communications Society.

[WGS+08] F. Wang, A. Ghosh, C. Sankaran, P. J. Fleming, F. Hsieh, and S. J. Benes. Mobile wimax systems: performance and evolution. *IEEE Communications Magazine*, 46(10):41–49, 2008.

[Wil99] F. Williams. Fourth generation mobile. In *Proceedings of the 4^{th} ACTS Mobile Communications Summit 1999 (AMOS'99)* [zz:99].

[Wil09] S. Wildstrom. AT&T's Stephenson: The road to 4G, 2009.

[WMH02] G. Wu, M. Mizuno, and P.J.M. Havinga. MIRAI architecture for heterogeneous network. *IEEE Communications Magazine*, 40(2):126–134, Feb 2002.

[XJH07] L. Xia, L.-G. Jiang, and C. He. A novel fuzzy logic vertical handoff algorithm with aid of differential prediction and pre-decision method. In *Proceedings of the 2007 IEEE International Conference on Communications (ICC'07)*, pages 5665–5670, Glasgow, Scotland, June 2007.

[XT04] A. E. Xhafa and O. K. Tonguz. Dynamic priority queueing of handover calls in wireless networks: an analytical framework. *IEEE Journal on Selected Areas in Communications*, 22(5):904–916, 2004.

[YGQD07] K. Yang, I. Gondal, B. Qiu, and L. S. Dooley. Combined SINR based vertical handoff algorithm for next generation heterogeneous wireless networks. In *Proceedings of the 2007 IEEE Global Telecommunications Conference (GLOBECOM'07)*, pages 4483–4487, Washington, DC, USA, November 2007.

[YMŞ08] X. Yan, N. Mani, and Y. A. Şekercioğlu. A traveling distance prediction based method to minimize unnecessary handovers from cellular networks to WLANs. *IEEE Communications Letters*, 12(1):14–16, January 2008.

[You79] W. R. Young. Advanced mobile phone service: introduction, background, and objectives. *Bell System Technical Journal*, 58:1–14, 1979.

[YPMM01] M. Ylianttila, M. Pande, J. Makela, and P. Mahonen. Optimization scheme for mobile users performing vertical handoffs between IEEE 802.11 and GPRS/EDGE networks. In *Proceedings of the IEEE Global Telecommunications Conference GLOBECOM'01* [IEE01].

[YŞM08] X. Yan, Y. A. Şekercioğlu, and N. Mani. A method for minimizing unnecessary handovers in heterogeneous wireless networks. In *Proceedings of the 2008 International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'08)*, pages 1–5, Newport Beach, CA, USA, June 2008.

[YŞN] X. Yan, Y. A. Şekercioğlu, and S. Narayanan. A probability based handover triggering condition estimation method for wlan usage optimization (submitted). *Wireless Communications and Mobile Computing*.

[Zha04] W. Zhang. Handover decision using fuzzy MADM in heterogeneous networks. In *Proceedings of the 2004 IEEE Wireless Communications and Networking Conference (WCNC'04)*, pages 653–658, Atlanta, Georgia, USA, March 2004.

[ZK03] T. Zahariadis and D. Kazakos. (R)evolution toward 4G mobile communication systems. *IEEE Wireless Communications*, 10(4):6–7, 2003.

[ZL05] A. H. Zahran and B. Liang. Performance evaluation framework for vertical handoff algorithms in heterogeneous networks. In *Proceedings of the 2005 IEEE International Conference on Communications (ICC'05)*, pages 173–178, Seoul, Korea, May 2005.

[ZL07] W.-H. Zhu and T. Lamarche. Velocity estimation by using position and acceleration sensors. *IEEE Transactions on Industrial Electronics*, 54(5):2706–2715, October 2007.

[ZLS06] A. H. Zahran, B. Liang, and A. Saleh. Signal threshold adaptation for vertical handoff in heterogeneous wireless networks. *Mobile Networks and Applications*, 11(4):625–640, August 2006.

[ZM04] F. Zhu and J. McNair. Optimizations for vertical handoff decision algorithms. In *Proceedings of the 2004 IEEE Wireless Communications and Networking Conference (WCNC'04)*, pages 867–872, Atlanta, Georgia, USA, March 2004.

[ZM06] F. Zhu and J. McNair. Multiservice vertical handoff decision algorithms. *EURASIP Journal on Wireless Communications and Networking*, 2006(2):52, April 2006.

[ZPK03] S. Zvanovec, P. Pechac, and M. Klepal. Wireless LAN networks design: Site survey or propagation modeling? *Radio Engineering Prague*, 12(4):42–49, 2003.

[zz:99] *Proceedings of the 4^{th} ACTS Mobile Communications Summit 1999 (AMOS'99)*, Sorrento, Italy, 1999.