Network Controlled Vertical Handover for Heterogeneous Network

Prashant Gami

[redacted]

Supervisor: Dr Kumbesan Sandrasegaran

August 2008

A thesis submitted in part fulfillment of the degree of
Master of Engineering by Research at
The Faculty of Engineering
University of Technology, Sydney

Certificate of Originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

Production Note:
Signature removed prior to publication.

Acknowledgement

This is a great opportunity for me to acknowledge kind support and encouragement I have received from all the friends, family, staff and colleagues. I would like to start with my supervisor Dr. Kumbesan Sandrasegaran (Sandy) without his inspiration and support at academic and personal level I would have not been able to complete this degree. I came in touch with him during the classes of wireless networking in my first semester of Master of Engineering Studies at UTS in 2006. He has been inspiration for me since the first day of the university. I have discussed so many things with him and every advice from him has given me a breakthrough during my studies and my career. His way of analyzing the things and planning in advance has always fascinated me. I have received some lifetime lessons along with academic knowledge from him. Since my wireless networking classes I got engaged in academic assignments and projects with Sandy. I took the project on handover in WLAN in NS2 during my second semester of Master of Engineering Studies which increased my interaction with Sandy. Finally I decided to undergo Master of Engineering by thesis under his guidance and I received effective guidance from hi throughout my degree. During my degree he has spent enormous amount of time in reviewing my research papers and guiding me in my work throughout the year. He has always been a guiding light at every step of my degree.

I would also like to thank my team mates Mo, Rachod, Leija and Bill who have helped me throughout my work and has provided useful guidance at all the stages. I would also like to thank my colleague Jamal who has been great company during my degree and staff members Phyllis, Dr. Prashanthi Hegare, and Dr. Robin Brown who has provided useful help during my degree.

I would also like to thank my parents Mr. Prabhulal Gami and Mrs. Pragna Gami, my brother Dr. Samir Gami and my bhabhi Mrs. Mona Gami who have inspired me to study further. I would also like to thank my friends Miral, Amit, Dipabhabhi, Janak, Nirav, Krunal, Brijeshbhai and all who has been great support during the study and has taken care of all my schedules.

Table of Content

List of Figures

Abstract

Future wireless networks are expected to be a convergence of different wireless networks based on single Radio Access Technology. The aim is to provide seamless services to users anytime from anywhere in the world using proper devices.

In order to achieve this goal, seamless transition from one network to another with acceptable quality of service is required. Handover from one network to another is known as inter network or vertical handover.

Currently known Mobile Controlled Vertical Handover schemes have some limitations. Mobile Station is limited with power and has very limited information about networks. Due to these limitations, Mobile Controlled Vertical Handover (MCVHO) can not achieve high performance in terms of handover delay and handover decision.

In this thesis, a new Network Controlled Vertical Handover (NCVHO) scheme is proposed. Proposed NCVHO scheme uses information which is available only to network along with information available to mobile node to enhance handover decision and reduce total number of handover and call dropping.

NCVHO scheme also reduces the signaling on the wireless part of the network which reduces the usage of air interface. The scheme allows Point of Attachment (PoA) to make handover decision on behalf of mobile node which reduces signaling on the wireless part of the network. This change in a vertical handover scheme reduces signaling delay of the handover and achieves better handover delay performance.

NCVHO scheme is reduces signaling on the air interface and uses more information to make handover decision which can achieve higher performance for number of handovers and call dropping than MCVHO. The price is paid as increased processing delay at the Point of Attachment (PoA). Simulation in this thesis shows that NCVHO performance decreases at higher number of users in comparison with MCVHO.

A solution to this problem is also proposed in this thesis with some modification to NCVHO scheme, a new network controlled vertical handover scheme called distributed Network Controlled Vertical Handover (distNCVHO) is proposed. The scheme achieves better performance than NCVHO scheme in terms of handover delay, number of handovers and call dropping.

It is noticed in this thesis that more information available from the network can enhance handover decision which can in turn reduce total number of handovers with a large margin. Information such as network load, network properties such as coverage, available bandwidth can play a vital role in the handover decision making. Reduced usage of air interface can allow network to perform handover faster and reduce number of total handover attempts.

1 Introduction

1.1 *History of wireless networks*

1.1.1 Evolution of Cellular Networks

Cellular communication has grown enormously over the past two decades witnessing the evolution of the fourth generation (4G) cellular systems from the first generation (1G) systems of the 1980s (Figure 1.1). Analog transmission has been replaced with the advanced digital high-speed network today.

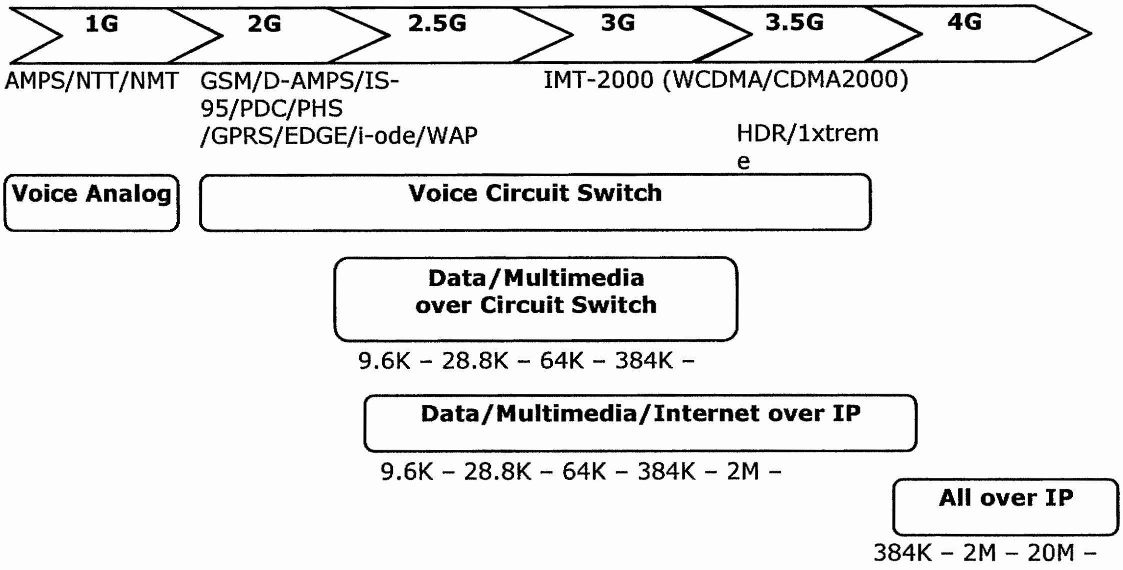


Figure 1.1 Evolution of Cellular Systems [2]

1G cellular communication systems, which were in use in the early 1980s, were analog networks such as Advanced Mobile Phone Service (AMPS), Total Access Communications System (TACS), and Nordic Mobile Telephone (NMT). The analog networks employed the Frequency Division Multiple Access (FDMA) method for allowing access by multiple users and techniques such as frequency reuse for maximizing frequency utilization [1].

Second generation (2G) cellular systems, introduced in the late 1980s were based on digital transmission techniques. 2G systems were superior to 1G systems in terms of capacity, quality, flexibility, security, and system complexity. 2G systems used Time Division Multiple Access (TDMA) or Code Division Multiple Access (CDMA) in combination with FDMA for sharing the spectrum more effectively. An example of a CDMA system developed by the CDMA Development Group (CDG) and the Telecommunications Industry Association (TIA) is the 2G system IS-95, and an example of a TDMA system, which was used in North America, is `the dual mode analog–digital 2G system called IS-136 (also known as Digital AMPS (DAMPS)). It was at this point in time that the 2G system called Global System for Mobile Communication (GSM) was developed by the European Telecommunications Standards Institute (ETSI). 2G systems were primarily designed for circuit switched transmission and voice services.

An intermediate stage called 2.5G, which was during the transition from 2G to 3G, introduced some enhancements to 2G to increase data rate and thereby upgrade 2G voice services, so as to provide multimedia services. This upgrade was done as a phased approach. The 2G HSCSD (high-speed circuit switched data) technique was used with a maximum transmission rate of 57.6Kbps in accordance with the GSM/HSCSD specification. In addition, General Packet Radio Service (GPRS) was also used to provide a higher data rate using packet switching, up to 115Kbps, leading to Enhanced Data rates for GSM Evolution (EDGE) which supports data rate up to 384Kbps [1].

3G systems were designed with the objective of further improving the data rates. In 3G systems, GSM systems were enhanced to Universal Terrestrial Mobile Services (UMTS), which is based on the Wideband CDMA (W-CDMA) technology led by the Third Generation Partnership Project (3GPP) in Europe. Meanwhile, the CDMA was also upgraded to a 3G system known as CDMA2000 by the Third Generation Partnership Project 2 (3GPP2) in North America. The aim of 3G systems was to provide global coverage under one technology. Although 3G are supposed to provide a data rate up to 2 Mbps theoretically, a user can achieve a maximum data rate of only 384Kbps, in the practical setting. 3G systems have been used in many countries including Australia.

Another type of a popular wireless network is Wireless Local Area Network (WLAN), which is widely used to access data services such as Internet, web browsing, and video streaming.

1.1.2 Evolution of Wireless LANs

The immense growth experienced by WLAN in the past decade is mainly because a WLAN allows a user to access high-speed data rate in a small coverage area known as a hotspot and uses an unlicensed frequency band. The IEEE 802.11 Working Group (WG), with the motive of enhancing WLAN standard to provide better security, Quality of Service (QoS), and coexistence with other technologies, is improving dominant 802.1x wireless standards. The evolution of WLAN is shown in Figure 1.2.

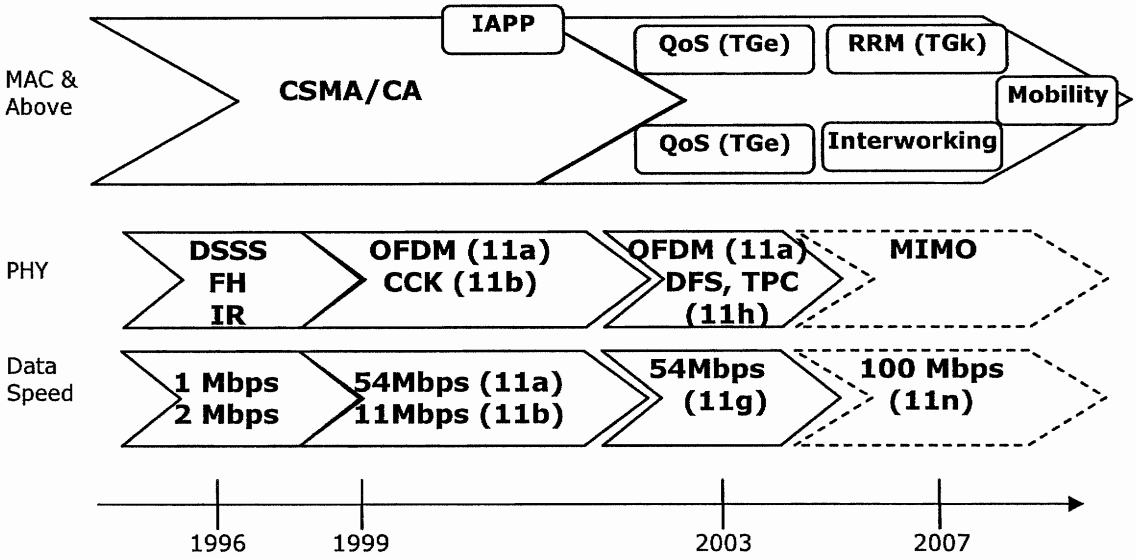


Figure 1.2 Evolution of Wireless LAN [3]

The International association IEEE 802.11 WG was formed in 1989 to develop WLAN specification. Seven years later, the first WLAN standard 802.11 was released, which specified three physical layer (PHY) technologies and a unified medium-access control (MAC) technology. This standard supported a data rate of 1 Mbps and 2 Mbps over wireless medium. The MAC layer technology of WLAN is based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), which includes mandatory

Distributed Coordination Function (DCF) and optional Point Coordination Function (PCF), whereas the PHY of this standard is based on Directed Sequence Spread Spectrum (DSSS) or Frequency Hopping Spread Spectrum (FHSS) technology.

In 1999, the WLAN standards 802.11b and 802.11a, which use the 2.4 GHz and 5 GHz frequency bands respectively, were released. 802.11b uses DSSS as PHY technology, similar to 802.11, and supports speed up to 11Mbps, whereas 802.11a uses Orthogonal Frequency Division Multiple Access (OFDM) as PHY technology and supports speed up to 54 Mbps.

In 2003, the 802.11g standard was released, which uses the 2.4 GHz frequency band. This standard supports both DSSS and OFDM as PHY technologies and CSMA/CA as a MAC technology, and provides a data rate of up to 54Mbps. The newly released WLAN standard 802.11n uses the Multiple Input Multiple Output (MIMO) technology and can provide a data rate that is twice than 802.11b and 802.11g [7].

Nevertheless, there exists a major limitation on the coverage capacity of WLANs. Higher data rates require a broader frequency band, which is only available at higher frequencies. Therefore, as the propagation path loss is proportional to the frequency band, increased path loss will occur at higher frequencies band.

The growth in user demand and diversity of services has led the WLAN WG and 3GPP to draft new strategies combining the advantages of WLAN and cellular technologies. The standard specifies aspects such as common charging, authentication and handover mechanisms [7]. This convergence of cellular networks and WLAN is providing a path for evolution to 4G cellular networks.

1.2 4G

4G cellular technologies are intended to complement and replace the current 3G systems. This technology aims to provide seamless service across all the networks around the world, support high-speed multimedia services and access to high volume of information including data, pictures, and video. The common layer interconnecting all these technologies is the Internet protocol (IP) [4]. The following are some of the major advantages of 4G technology:

- It will support high bandwidth, higher data rate up to 100 Mbps, and smoother transition from one network to another [4];
- It will include various networks including public and private networks;
- It will interoperate with 2G and 3G networks with digital broadcasting;
- It will have all IP-based wireless internet access.

Currently, 3G wireless networks are used in many countries and 4G cellular networks are in the planning phase. A number of interpretations exist with regard to the 4G systems. The transition to 4G is not a change in an interface technology like the previous transitions, but it is an integration of the current technologies and the convergence of indoor technologies such as WLAN 802.x and Bluetooth and outdoor technologies such as GSM/EDGE, W-CDMA, CDMA, CDMA2000, and Satellite Networks [4]. The greatest challenge of 4G networks is to merge all these technologies. The integration of cellular networks will result in wireless overlay networks, where many networks are overlapped, one above another in a same geographical area. The important features of 4G are detailed in Reference [4] and [5].

1.3 Cellular Networks

Generally, a big geographical area served by a network is divided into cells. Users can access services using a mobile node known as Mobile Station (MS). Each cell is served via a network node known as Base Station (BS). Users are served by a BS in a cell whereas an MS acts as a client, accessing the services of a network. The BS is the only node that establishes communication between the wired and wireless parts of the network [11]. Communication by a BS with the network is through a wired or wireless link and communication with an MS is through a radio link. Thus, a user can access services continuously while on a move by means of switching cells in the network (Figure 1.3).

All the calls in the network are switched by a network node known as Mobile Switching Center (MSC), which performs switching of calls over a large service area comprising many cells. Hence, the MSC is also known as Mobile Telephone Switching Office (MTSO) [11]. The MS communicates with the BS using a duplex channel; the channel from a BS to MS is known as forward channel, and the channel from an MS to BS is known as reverse channel. The interference between two adjacent channels is known as Co-Channel Interference (CCI) [11].

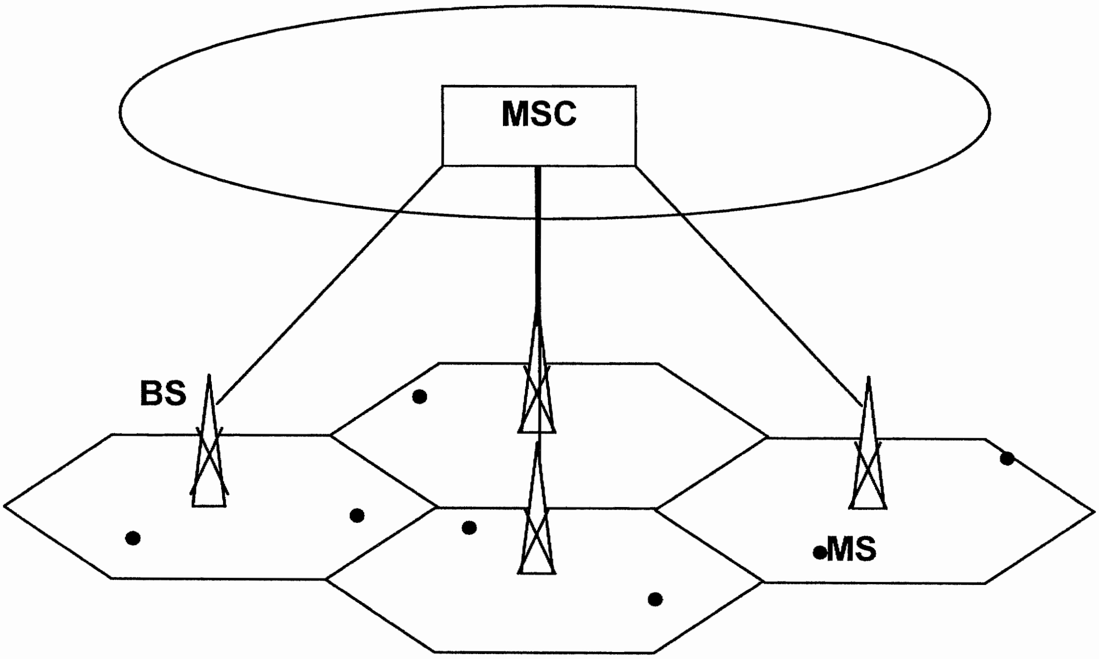


Figure 1.3 Cellular Network [6]

The phases involved in the planning of a cellular network includes assessment of traffic density, determination of cell size, choice of omni-directional or sectored antenna, antenna directions, selection of BS sites to cover the required area, allocation of frequencies and selection of power control parameters, and handover parameters [3]. Effective planning and implementation of all these phases are necessary to ensure good service.

1.4 Handover in Cellular Networks

As the MS moves toward an edge of that cell, the Received Signal Strength (RSS) from the BS starts decreasing as the path loss and other environmental losses increase. This decrease in RSS results in a drop in the level of service experienced by the user. This is shown in Figure 1.4: as the MS is moving away from BS1, the RSS from BS1 is decreasing. At a point when RSS from BS1 decreases below an acceptable level, MS starts looking for other BS in the vicinity with a better RSS. Due to signal fluctuation in this region, unnecessary handoff may occur which is known as ping-pong effect. To avoid this, a handover margin is introduced. At the point when the RSS from BS2 exceeds that received from BS1 by a handover margin, the MS changes its radio connection to BS2. This procedure of an MS changing a radio connection from one BS to another is known as handover [4].

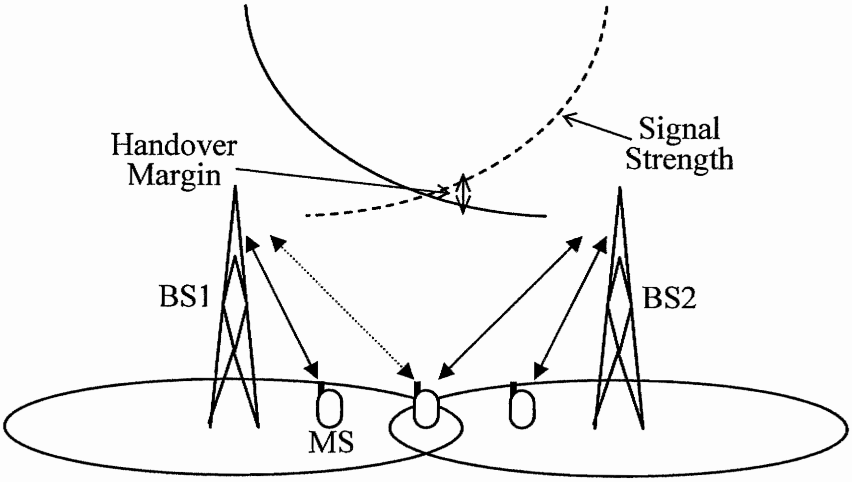


Figure 1.4 Handover in Cellular Network [2]

1.5 Handover in Heterogeneous Networks

The next generation of wireless networks will be a convergence of many wireless networks using a common IP-based core network. The aim is to provide seamless services to the users. This convergence will allow users to access services anytime, anywhere, from any network using appropriate devices. To achieve this goal, it is essential to ensure seamless handover mechanisms both within the network and among the contributing networks. In heterogeneous networks, it is possible that all the networks can be present at a location, laid over one another, which is called wireless overlay networks.

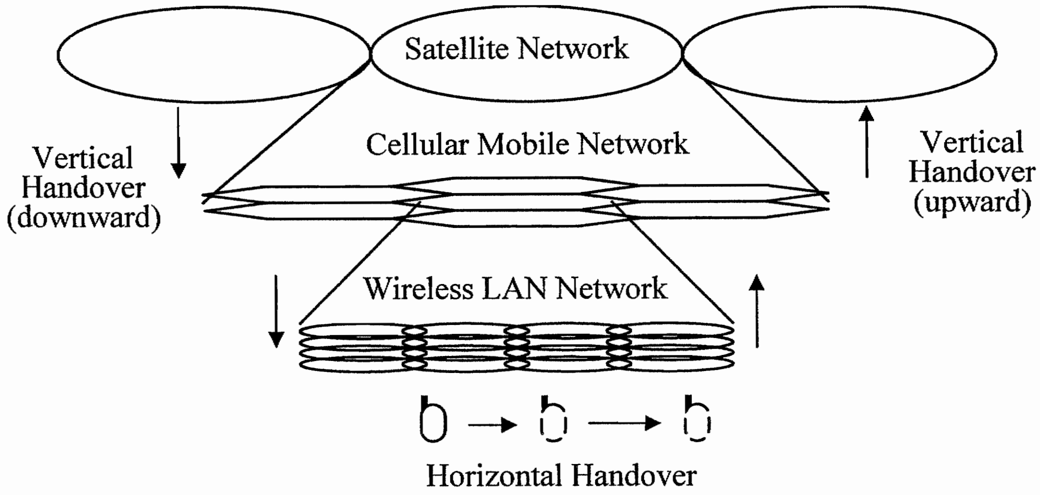


Figure 1.5 Wireless overlay network modified from [5]

Figure 1.5 shows a wireless overlay network in which a satellite network is at the top, covering a wide area. The satellite network is laid over a number of cellular mobile networks each having a large coverage area usually of a country or province. The smallest unit of coverage is a WLAN, which covers an office building or a shopping mall. At a point within the WLAN coverage, all the networks are present. In a case when these networks are connected with each other, they allow users to switch between different networks seamlessly. For example, WLAN is the most suited option for high data rate such as video streaming or downloading, whereas cellular mobile network is the best option for more reliable services such as voice. Satellite network offers better services such as Global Positioning System (GPS).

In homogenous cellular networks, handover occurs within the same network. In heterogeneous networks, the Points of Attachment (PoA, which can be either a BS or an access point) between which handover is performed can belong to either the same network or different networks. If the PoA belongs to the same network, then the handover is known as intra-system or horizontal handover, whereas if the PoA belongs to different networks it is known as inter-system or vertical handover. If a vertical handover is performed from an upper network to a lower network then it is known as downward vertical handover and if it is performed in the reverse direction, then it is known as upward vertical handover.

In the case of a horizontal handover, a decision is taken based on conventional criteria such as RSS and Signal-to-Interference Ratio (SIR). However, in the case of a vertical handover, the handover process becomes more complex because more than one network is involved. Conventional criteria are insufficient to make a handover decision in this case. More information such as service requirements, network capacity, available bandwidth, and user preference are also needed. Some of the decision methods for heterogeneous handover are complex methods in comparison to horizontal handover decision making such as Analytical Hierarchy Process [7], Decision Metric [8], Fuzzy Logic [5], Neural Network Techniques [5], and Cost Function [9], which use the earlier specified information.

Vertical handover is triggered mostly when certain service requirements are not fulfilled by the current network or when a user moves out of a network coverage. Some of the vertical handover requirements are minimizing handover delay, number of handovers, call-dropping probability, the use of air interface, considering user preference, and ping-pong effect [10]. These requirements are discussed in detail in Chapter 3.

As mentioned earlier, the complexities involved in a vertical handover are more than those in a horizontal handover mainly because more than one network is involved. As a result, the latency in performing a vertical handover is expected to be longer than in horizontal handover; this latency should however be minimized to provide seamless service.

A handover procedure can be controlled by either network or MS. If the handover procedure is controlled by the MS then such a handover is known as Mobile Controlled Handover (MCHO) and if it is controlled by the network then it is known as Network Controlled Handover (NCHO).

1.6 Problem Statement and Research Objective

The most popular approach to handle the complexities involved in a vertical handover in literatures is to have an MCVHO. In MCVHO, the mobile node makes all decisions about when to perform the handover and controls the entire handover procedure. However, this approach has the following limitations:

A. The Mobile node has limited information about the network conditions.
B. The Mobile node has limited battery power.
C. If the more information is collected from network, then it will increase signaling on the wireless part, which can cause greater delays in the handover procedure.

Taking these limitations of MCVHO into account, the research question addressed by this thesis is,

> *On the basis of the currently known methods of Mobile Controlled Vertical Handover, can one generate a more optimal Network Controlled Vertical Handover method that can use extra information available only to the network to achieve better performance in terms of delay, number of handovers, and call dropping?*

Aim:

The objective of this research was to develop a new Network Controlled Vertical Handover (NCVHO) scheme which can use information available to the mobile node as well as network to choose better target networks and to decrease the delay of vertical handover procedure.

Much of the literature focuses on the decision schemes for a vertical handover. However, most of these do not specify the source of information that was used to make the decision and try to achieve faster vertical handover by providing faster authentication. Some of the schemes presented in the literature are proactive authentication, reactive authentication, and takeover scheme [10]. However, there are two factors that are most essential in a vertical handover: enhancement of the signaling and target network selection. The latter factor is critically important because a wrong selection of the target network will result in a higher number of handovers and call dropping. If handover is controlled by the network, then a less amount of information will be passed on the wireless part of the network and the selection of target network can be enhanced using more information. NCVHO will also reduce the amount of signaling on the wireless part of the network. This thesis proposes a new NCVHO scheme that can reduce the signaling delay in the handover procedure. Better target network selection is also proposed using NCVHO, by achieving less call dropping and number of handovers. Improvements on the new NCVHO are proposed to reduce handover delay. A new handover scheme namely distNCVHO is proposed as a solution.

1.7 Thesis Outline

The thesis is organized as follows. Chapter 1 contains the introduction of the thesis. It describes history of cellular networks and WLAN, overview of 4G, handover in a cellular network, handover in a heterogeneous network and wireless overlay network. In the last part, problem statement and research objectives are listed.

Chapter 2 discusses Radio Resource Management (RRM) functions, future heterogeneous networks, Common Radio Resource Management (CRRM) and Authentication, Authorization and Accounting (AAA). Handover and its modeling techniques are also discussed in this chapter. Various techniques to model various parts of cellular networks and their importance are discussed in this chapter.

Chapter 3 provides study of a handover. Handover in a cellular network has been discussed in this chapter. This chapter includes causes of handover, complexities involved in handover, desirable features of handover, criteria used to make a handover decision. This chapter also includes conventional algorithms of handover and their limitations. Emerging handover algorithms and handover control frameworks are described in this chapter. Radio resource management is also discussed.

Chapter 4 proposes a new vertical handover scheme. This chapter describes MIPv6 based vertical handover scheme, MCVHO scheme and newly proposed NCVHO scheme. This chapter focuses on the performance measurement of a new scheme for handover delay with MIPv6 and MCVHO. The simulation results and analysis of handover delay involved in all three schemes has been presented in this chapter.

Chapter 5 presents a study on network selection schemes. This chapter also includes other performance analyses of the new scheme. The chapter presents a new simulation scenario which allows performance measurement of NCVHO in terms of number of total handovers, call blocking, call dropping etc.

Chapter 6 focuses on the limitation and enhancement of NCVHO. This chapter proposes solutions to those limitations and proposes Distributed Network Controlled Vertical Handover (distNCVHO) with some modification to the original NCVHO scheme proposed in previous chapters. Performance analysis of the new scheme is presented in this chapter.

Chapter 7 finally concludes the thesis.

1.8 Original Contribution

Following contributions are made in this thesis which are considered original,

Chapter 4 –

- A new vertical handover scheme NCVHO is proposed
- Performance Comparison of NCVHO with existing vertical handover algorithm for heterogeneous network for signaling cost and handover delay

Chapter 5 –

- Performance comparison of vertical handover decision algorithms with different parameters used to make handover decision
- Performance comparison of NCVHO and MCVHO in terms of number of total handovers, and call dropping

Chapter 6 –

- Distributed Network Controlled Vertical Handover (distNCVHO) as an Enhancement to the proposed NCVHO to decrease processing delay at Node B
- Performance comparison of distNCVHO and MCVHO

As a result of this work, an IET Letter is submitted which is accepted for publication in April 2008 issue of IET Letters,

P. P. Gami, K. Sandrasegaran, “A new network Controlled Vertical Handover”, *IET Letters*, 24th April 2008, Vol. 44 No. 9

2 Background

2.1 Radio Resource Management

Radio resources are valuable assets in the field of wireless communication. The applications supported by wireless networks are increasing day by day, which, in turn, increases the complexity in radio resource management. Efficient management of radio resources can increase the performance of the cellular system significantly. Radio resource management tasks are admission control, power control, congestion control, handover control, and packet scheduling. The ultimate goal of radio resource management is to achieve high capacity and throughput for a given QoS. It is possible to combine individual radio resource management tasks, e.g., channel assignment and handover can be combined [38] into a centralized RRM entity. The goals of the centralized radio resource management entity can compromise with the individual goals of these tasks to achieve better performance of the overall system [56]. The radio resource management tasks are described in the following sections.

2.1.1 Admission Control

Depending on available radio resources, the Radio Network Controller (RNC) will decide on whether to accept new calls into the network. This procedure is known as admission control. RNC will decide on the amount of resources needed to accept a new call using the amount of resources available with the network. If the amount of resources available in the network is less than needed by the call, then the call will be rejected. New calls and ongoing calls can be treated differently in the admission control by the RNC. The criteria used by different wireless networks for admission control may be different. For example, TDMA and FDMA based cellular system such as GSM are limited by the number of available channels, CDMA based systems such as UMTS are limited by the transmission power of the mobile user. In most networks, admission control is performed by defining thresholds for various network parameters.

2.1.2 Congestion Control

Congestion control is the procedure for handling congestion in a network. A number of different approaches exist for handling congestion. This RRM function is responsible for checking whether a network has reached its maximum load limit, and if the network has reached its maximum limit then bring it back to a stable state if it has not yet attained the same [1]. For example, as shown in the Figure 2.1, if a Cell A has reached its maximum load limit, it may handover users in the edges of the Cell A to a neighboring Cell B.

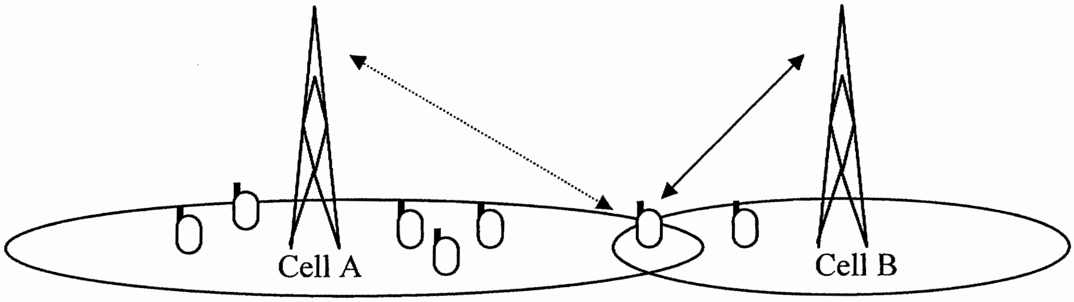


Figure 2.1 Handover as a Congestion Control Mechanism [3]

As shown in Figure 2.1, if a user in Cell A starts moving toward the edge then the transmit power of the Base Station in Cell A starts increasing. This creates high interference in other cells using the same frequency band. To avoid this interference, Node A will initiate handovers to neighboring cells for reducing the power requirements of the current cell.

2.1.3 Power Control

There are three types of power control in cellular networks: open loop power control, outer loop power control, and inner loop power control.

Open Loop Power Control

Open loop power control is used when a user accesses the network for the first time, and when it has not been allocated a dedicated channel [47]. As a dedicated channel has not been allocated, channel estimation for this channel is not possible. Nevertheless, the network needs to know some information to, for example, set initial transmission power. Therefore, the radio channel from an opposite direction i.e. downlink in an uplink communication case (user making call). In an uplink communication, User Equipment (UE) uses the pilot signal it received from a downlink common channel to set its initial transmission power. Given that the minimum required signal to interference plus noise ratio ($\mathrm{SIR}_{\mathrm{min}}$) is known to UE, it will then know the transmission power needed, which is calculated using the following equation [47].

$$P_T = L \cdot I_m \cdot SIR_{\min} = \frac{P_{T,pilot}}{P_{R,pilot}} \cdot I_m \cdot SIR_{\min} \qquad \text{... Eq. 2.1}$$

Where $P_{T,pilot}$ is the known transmission power of the pilot signal

$P_{R,pilot}$ is the known reception power of the pilot signal

I_m is Interference Power

SIR_{min} is the minimum SIR required for a reliable connection and

L is the propagation loss on the radio link.

Outer Loop Power Control

Outer loop power control is used to set the target SIR for the inner loop power control in the network. Outer loop power control decides the SIR target on the basis of some parameters, e.g., the Frame Error Rate (FER) [47]. The SIR target for each link varies because the quality of the radio link connection varies. The main objective of outer loop power control is to set SIR target so that the quality of the radio link can be maintained on each communication link.

Inner Loop Power Control

Inner loop power control is performed to match transmission power with the SIR target set by the outer loop power control. It is also known as closed loop power control [47]. It dynamically adjusts the transmit power of a mobile station so that rapid changes on the radio link as a result of path loss and fading can be handled while maintaining the communication quality. The BS (Node B in UMTS) sends Transmission Power Control (TPC) commands every time interval, which is used to inform UE to adjust its transmission power to achieve the SIR target. If the received SIR is higher that the target, then the power will be decreased; otherwise the power will be increased.

2.2 Future Heterogeneous Networks

Future wireless networks are expected to be based on a common IP-based core network through which a number of networks can interoperate. These networks can be based on different radio access technologies as shown in Figure 2.2.

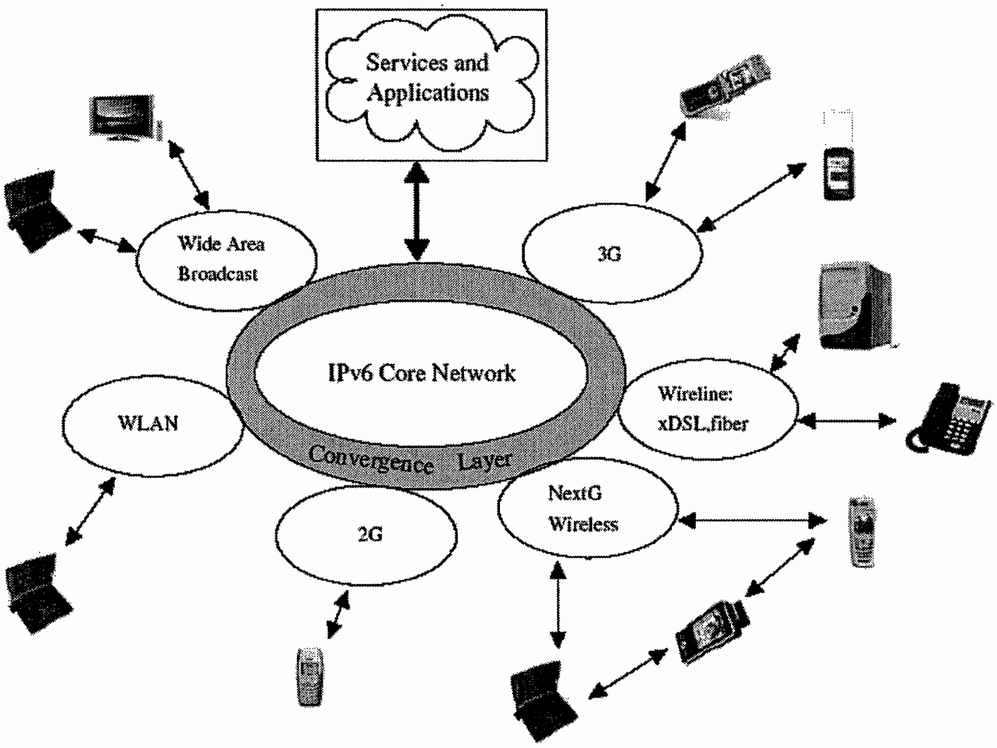


Figure 2.2 Future Heterogeneous Networks [13]

Figure 2.2 shows interoperation between different networks such as WLAN, 2G, 3G, NextG and WAN, connected to a common IP-based core network via a convergence layer. The network that comprises many networks with different (heterogeneous) access technologies may be referred to as a heterogeneous network. In such a heterogeneous network, the contributing networks and the available services and applications may be managed by different service providers, independent of each other. For this purpose, new entities such as Authentication, Authorization, and Accounting (AAA) and Common Radio Resource Management (CRRM), which support faster authentication and better Radio Resource Management (RRM) across different networks, are required to ensure seamless and efficient interoperation. These entities are detailed in the following sections.

2.2.1 CRRM

A heterogeneous network comprises many networks with different Radio Access Technologies (RATs). Individual networks that contribute to the formation of a heterogeneous network are based on single RAT, and each one is better suited than another for a particular type of service. One of the challenges of a heterogeneous network is to associate users to the most suitable network, given the user preferences and the service type. Most current RRM techniques have been researched and developed to implement on a network which single RAT only and are not designed for a heterogeneous network. To achieve an effective radio resource management across different networks, a set of more advanced RRM techniques to take into account the incorporations between different networks with different RATs need to be proposed. The advanced RRM is often referred to as a Common Radio Resource Management (CRRM) [10]. CRRM is based on a two-tier model of the RRM. The lower tier composes of conventional RRM entity, which manages the radio resources within an individual network, whereas the upper tier comprises an CRRM entity, which manages radio resources between all networks. This two-tier architecture is shown in Figure 2.3.

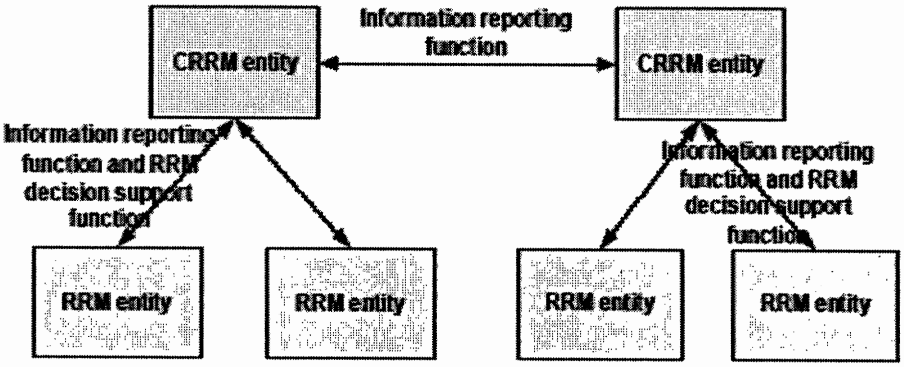


Figure 2.3 Common Radio Resource Management Architecture [10]

Figure 2.3 shows a conceptual block diagram of CRRM. Each RRM entity is responsible for monitoring and managing its respective network. Also, it has to report information to the CRRM. The CRRM uses this information together with information received from other RRM entities to either assist the RRM entities making a decision, or CRRM entity itself may make a decision for individual network. This is still an open research topic. The aim of CRRM is to connect users to the most appropriate network every time. This is achieved using two functions of CRRM: information reporting and RRM decision support.

The information reporting function, as the name suggests, allows RRM entities to report important information about network conditions to the CRRM entity, which gathers information such as current congestion in the network, available bandwidth, power requirements of the network cells, admission criteria for that instance in the cell, and coverage capacity of the cell. This information can be very useful in managing radio resources across networks and allows users to access services in a better way with low call blocking and dropping.

The CRRM entity collects information time to time from each RRM. The frequency of information update may vary depending on the type and purpose of the information provided by the RRM entities. For example, such information as network congestion, available capacity, and user position need to be reported from RRM entities to their corresponding CRRM entity on a fast basis, for example, every 10 ms. On the other hand, such information as policy translation, initial RAT selection, and admission control related parameters may need a lower interaction between RRM and the CRRM entity.

2.2.2 AAA

Authentication is a process of verifying claimed attributes, authorization is a process of checking whether a particular right should be given to the entity, and accounting is a process of collecting information for billing purposes. These are critically important functions in any wireless networks. Each wireless network uses its own policies for AAA. A user has to be authenticated in order to access the services of a particular network. When a user changes his or her connection from one network to another, he or she has to be authenticated again with the new network. According to the MIPv6-based authentication scheme, this authentication is done via the Home Agent (HA) of an MS. The HA of an MS is usually connected to the current network of the MS through a public network such as Internet. If the authentication is performed via the HA, then the delay may be high because the signals have to travel through the public network, and this can lead to higher handover latency.

To decrease the latency associated with the handover procedure, a AAA agent architecture has been proposed in [15]. The AAA agent architecture consists of AAA agents, which can act as either as a AAA server or as a AAA broker. The former stores all the information needed for authenticating the user in a network, whereas the latter negotiates the authentication with the target networks and authenticates an MS to the network using fast authentication schemes. The AAA agents create a trust network which allows networks to trust each other via a common AAA architecture. These trust relations allow faster authentication of users who are already authenticated by trusted networks. The information provided by the trusted network can be accepted by the current network. This trust relation allows faster authentication of MS during handover and as a result speeds up the handover process.

The trust relation between the networks can be established using various policies as suggested in various literatures. The basic idea of the trust relation is that, at the first time a user connecting to a trusted network, a full authentication process is performed. After completion of the authentication process, the user is given a trust authentication key to indicate to other networks that this particular user has been authenticated with a trusted network. This key will be used in another authentication process with another network which has trust relationship with the previous network. Two examples of the AAA architectures proposed by IETF are RADIUS [15] and Diameter [15], respectively. These architectures allow faster authentication of an MS when they move from one network to another. The handover schemes can use such architectures and thereby decrease handover latency.

2.3 Cellular System Modeling

2.3.1 Network Modeling of Handover Study

Modeling a network for the study of handover mechanism can be done in several ways depending on what aspect of handover that is being studied. For instance, if the system aims to study a proposed handover scheme with its signaling and delays, a detailed network modeling is required. However, if the system is used for studying certain authentication schemes for handover, modeling can be simplified because the scheme is affected only by the relationship between networks and the number of networks or nodes. The following aspects are to be taken into consideration in the case of a detailed network modeling.

1. **Network Architecture Description**: This describes the structure of network being modeled, i.e. the components of networks, the way the components are connected with each other and with outside networks. This also includes whether there are wired and wireless layers in the network as well as the overlaid position of the networks. In general, different types of modeling can be used to model this; the most popular techniques are graph and tree.

2. **Coverage Specification:** This specifies a definition of how many PoAs are available, what their specifications are, the antenna specification which can be omni-directional or sectored, their sensitivity, power needs, and positioning. Depending on the design of the network, the cell coverage has shape and properties. In most networks the cells are modeled as hexagonal or circular; some networks are modeled using a grid structure as well.

3. **Capacity Specification:** This defines the capacity of the network which can be based on the available bandwidth, power needs, and noise present in a system. A number of factors can limit the capacity of the system. In the case of GSM, the system is limited with respect to many factors such as the bandwidth and the number of time slots

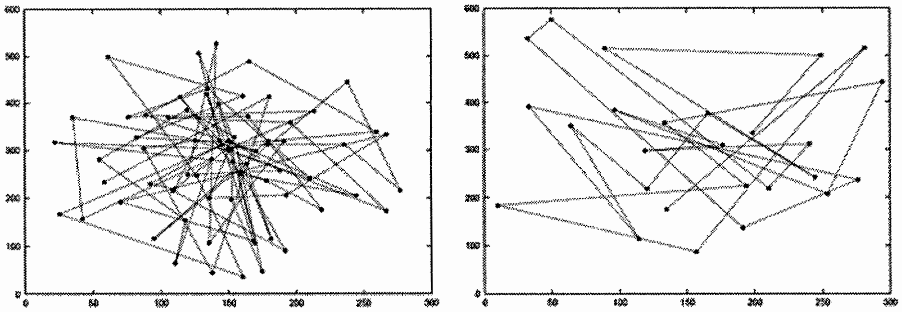
available, whereas a UMTS system is limited with respect to much more factors such as the uplink transmission power of UE, availability of orthogonal codes, cell configuration, use of scrambling codes, and interference level.

4. Device Description: This aspect of network modeling includes the capability of the device. In a heterogeneous network modeling, some devices may be more suitable for high-bandwidth network like WLAN whereas others may be more suitable for high-mobility networks like UMTS. The parameters of the device need to be defined to decide what power it will be transmitting, what services it can access, and how it will communicate with the networks.

As mentioned above, detailing of each part is decided on the basis of what aspect of handover is studied. In some cases, some network parameters are taken as ideal, whereas in some other cases discrete values are taken for some parameters. On the basis of our interest of study, different techniques can be used.

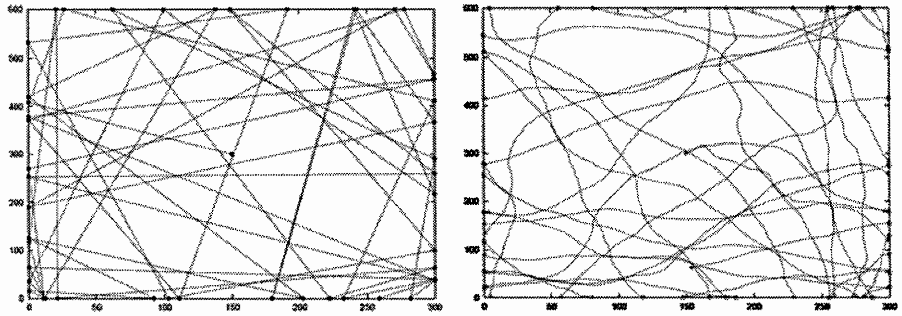
2.3.2 User Mobility Modeling

In a wireless network, the movement patterns of users play a vital role in managing services. In a real world scenario of wireless network, users move at speeds varying from 0 to 120 kmph. The modeling of user movement is very important for simulating a wireless system. In most cases, user movement is modeled based on some mathematical models. Depending on the different characteristics of the system such as indoor, outdoor, low coverage, and high coverage, different models are used. Some well-known user movement modeling techniques are include, for example, Brownian model, random walk model, random waypoint model, random direction model, boundless simulation model, Markovian model, incremental model, and mobility vector model [10]. Figure 2.4 shows four different user behavior models, namely the random walk model, random waypoint model, random direction model, and boundless simulation model. User mobility is modeled based on many parameters such as velocity (stationary or moving), mobility frequency (low or high), and situation (working or residential).



Random Walk Mobility Model

Random Waypoint Mobility Model



Random Direction Mobility Model

Boundless Simulation Area Mobility Model

Figure 2.4 User Behavior models [10]

User behavior modeling can also be categorized into two categories. In the first category, users are moved explicitly to generate vertical handover, whereas in the second category the user behavior model along with the traffic model implicitly triggers a handover. In this model, a user's behavior is often modeled using some probability function. For example in [10], a probability function named multidimensional probability density function is used to model user movement in indoor/outdoor etc. environment. Overall, the aim of the target system is the main factor in the selection of user mobility model in the system.

2.3.3 Radio Propagation

In a wireless system, the transmitted signal over a radio path is altered by factors such as path loss and multipath fading. A signal can be refracted, diffracted, or scattered in free space. Because of the complexity and the unpredictable properties of the environment, it is almost impossible to obtain exact knowledge of radio channel attenuation. For this reason, simplified mathematical models have to be used in a simulation. Radio propagation can be modeled on the basis of scale, characteristics of the environment, and the application area. Radio propagation can be modeled as either large scale or small scale. Large-scale models includes free-space path-loss models such as log distance path-loss model and log normal shadowing models, outdoor models such as 2-ray ground reflection model and diffraction model for hilly terrain, and indoor models such as log-normal for Line of Sight (LOS) down corridor and log-normal shadowing for no LOS and partition.

Path loss is a large-scale fading which depends on the distance between a transmitter and a receiver. A detailed model for free-space path loss was proposed by Okamura [11]. This model was then used by Hata [11] to provide an equation for modeling path loss in a simulation in 1960. However, the model was developed for a specific frequency range so that it cannot be used to model path loss in modern wireless systems. The model was, therefore, improved and changed to model path loss for the modern radio propagation system, and the new equation is named COST-231, which is given below.

$$Lu=46.3+33.9*\log(f)-13.82*\log(hb)-a(hm)+(44.9*6.55\cdot\log(hb))*\log(d) \qquad \text{Eq.2.2}$$

$$a(hm)=(1.1*\log(f)-0.7)*(hm)-(1.56\cdot\log(f)-0.8), \qquad \text{Eq.2.3}$$

where Lu is a propagation loss in dB;
f is the frequency of the transmission in MHz;
hb is the height of base station or transmitter in meters (30–200m);
hm is the height of the mobile or receiver in meters (1–10m);
d is the distance between the receiver and the transmitter in kilometers;
and a(hm) is the mobile antenna correction factor.

The signal strength of a signal affected by path loss is shown in Figure 2.5. For the simulation following values are used in Eq. 2.2 and Eq. 2.3.

f	is the frequency of the transmission = 1900MHz;
hb	is the height of base station or transmitter = 20m
hm	is the height of the mobile or receiver = 1.7m
d	is the distance between the receiver and the transmitter = 0.01 to 10 km

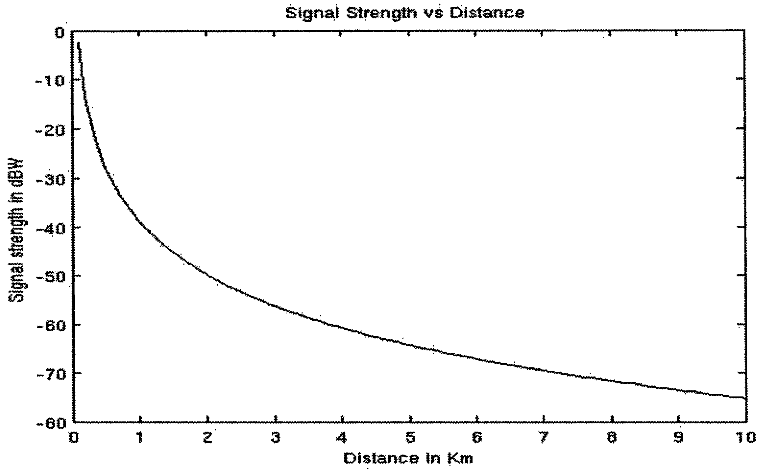


Figure 2.5: Signal Strength with Free-Space Path Loss

Multipath fading occurs due to the fact that radio signal takes multiple paths traversing from the transmitter to receiver in a wireless medium. The signal arrives at the receiver in different phases, delays, and attenuation which cause reduction in the signal strength. This phenomenon is known as "Rayleigh fading" or "fast fading." A signal received at the receiver is the sum of all signals traveled through different paths resulting in distortion of received signal amplitude. This received signal amplitude can be represented by following equation [17]:

$$h(t) = E_0 \cdot \sum_{n=1}^{N} C_n \cdot \exp(j(2 \cdot \pi \cdot f_d \cdot t \cdot \cos\alpha_n + \phi_n)) \qquad \text{Eq. 2.4}$$

where,

E_0 is the scale constant;

C_n is the random amplitude;

j is equal to $\sqrt{-1}$;

α_n is the angle of an incoming wave;

Φn is the initial phase associated with nth propagation; and

f_d is Doppler frequency.

Using the above equation, Rayleigh fading can be simulated, which is represented in Figure 2.6. The graph is generated using filter available in MATLAB with the sampling time 1ms and maximum Doppler shift of 10Hz.

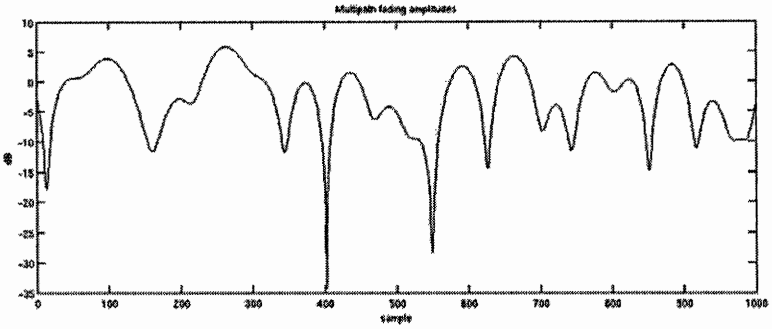


Figure 2.6 Multipath Fading Amplitude

2.3.4 Traffic Modeling

Traffic models in a wireless network are one of the most important aspects to be studied in handover simulation. In future networks, traffic models may include different types of traffic such as video streaming and web browning besides only voice traffic as considered in traditional mobile communication systems. These diversities in traffic offer different load on channel. For example, web browsing services require low data rate and non-continuous traffic flow in one direction (typically downlink). On the other hand, video streaming incurs heavy and continuous traffic in the downlink direction. For voice traffic, the traffic requirement is constant for call duration in both directions (uplink and downlink). An efficient traffic model for future networks should take all these diversities into account. Typical channel utilization for a mixed traffic is shown in Figure 2.7.

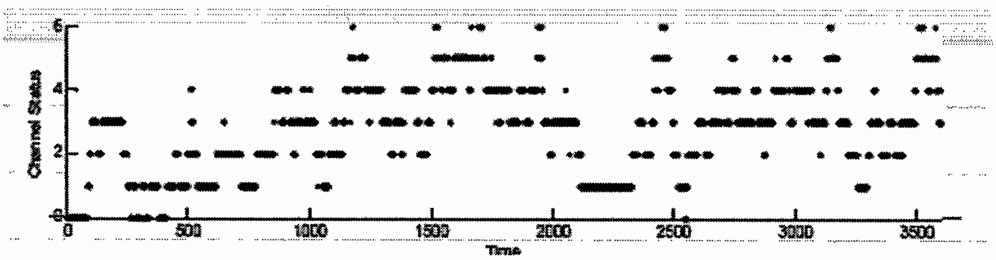


Figure 2.7 Channel Traffic

Figure 2.7 shows number of active users on a channel at a given time. Besides being affected by the type of services, traffic in a radio network is also affected by mobility-related factors. Some of the important factors are handover rate, call frequency, cell residency time, and channel utilization. Handover rate refers to the number of handovers that are performed on an average during service delivery. Call frequency refers to the number of mean calls that arrive in a network within a given period, typically one hour. Cell residency time is the mean time for which a user stays in a cell. Channel utilization is a factor indicating the resource usage of an application, which is mostly affected by the type of application. Effective modeling of traffic in a network must be provided to ensure effective performance evaluation of a proposed scheme on such network.

2.3.5 Power Control Modeling

Power control in a wireless network is very important for maintaining the quality of service (QoS). Any increase in transmission power will increase interference and degrade the QoS experienced by other users. Therefore, each wireless network has its own power control mechanism. In a heterogeneous network, power control is modeled individually for each connection in each network in accordance with their respective specification given in the standard.

2.3.6 Performance Evaluation System Modeling

A performance evaluation system is composed of various modules. In a vertical handover performance evaluation, this system includes modules such as user session generation, radio wave propagation, power control, and handover scheme in addition to other supportive parts that play roles such as joining, monitoring, and managing all these individual modules. The system allows data gathering, analyzing, and graph or report generation to illustrate the performance of the proposed handover scheme. The design of these parts of the system could be in detail or in abstract form depending on their usefulness in achieving the target. Because of the nonavailability of the actual network specification, modeling of a heterogeneous network includes many predictions and assumptions on the basis of those discussed in the literature and standard bodies. The design phase includes the detailed analysis of a long range of past data to allow mathematical modeling of the proposed system. The entire system can be designed using top-down or a bottom-up approach. In a top down approach, initially system is design in an abstract way and then sub parts of the system are expanded as needed. In case of a bottom-up approach, subsystems are first developed individually and then joined together to form a target system. The top-down approach is preferred than the bottom-up approach because it allows more optimization in the overall system. A performance evaluation system is shown in Figure 2.7.

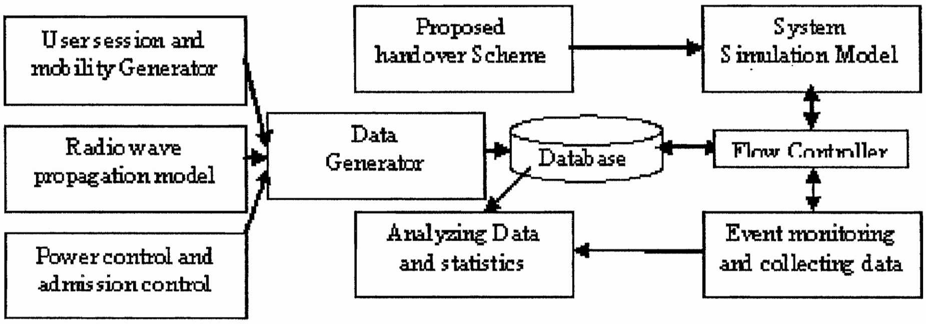


Figure 2.8 Performance Evaluation System [11]

The block diagram shown in Figure 2.8 is a performance evaluation simulation framework of a target system. This system simulates the model of a proposed handover scheme. To accomplish this, other RRM mechanisms and system modules such as user session, mobility generator, radio wave propagation, and power control are also simulated. These parts generate data which is collected in a database by a data management module. The flow controller keeps track of the execution of the flow of various parts of the system and triggers appropriate events in the system. It calls routines whenever needed and allows data collection at the end of each interval. The data needed for the analysis are collected by the event monitoring and data collection module. The analyzing and data statistics module analyzes the data generated during simulation by applying various statistical functions on generated data and generates either graphs or reports. Each module should be validated before finalizing the system. The efficiency of the system depends on effective implementation of individual parts and successful integration. Some of the most popular ways to represent the generated data are in the form of graphs, pie charts, and tables. The target system specified in the Figure 2.7 is a general model of any handover simulation which is performed to generate data and to analyze performance. This system is modified and customized to achieve different performance measurement goals. The evaluation of the performance of a handover process can be done with different aspects. These aspects are based on the cost involved in performing a handover. These costs are described in the following paragraphs.

Database access cost: This metric is generated to calculate time and space overhead to access the database to make a handover decision. For accessing data from the database, there is an access latency attached with it, and we need space to store all these data which occupies system memory and eventually the system becomes slow. Therefore, we need to optimize designing and the size of the metric carefully.

Signaling cost: This is the metric generated to calculate signaling overhead on the wireless and the wired part of the network. This metric can be used to calculate the actual handover latency on the basis of the amount of signals and the signaling delay on both parts of the network. The signals include paging signals, authentication signaling, and handover signaling.

Time cost: This is the metric that keeps track of all the delays in various parts of handover. These delays can be authentication delay, registration delay, packet delay, and so on. It focuses on the time taken by the system to generate performance parameters to complete a task involved in a handover such as authentication.

Based of these costs, performance evaluation can be done in terms of handover delay, usage of air interface, user throughput after performing handover, probability of call dropping, number of total handovers, QoS, user satisfaction, and minimizing ping-pong effect.

As discussed in this section, modeling a performance evaluation system for handover is quite complex. The system needs to take care of many designing issues to generate valid and useful data in an efficient and correct way. The system considerations show how the system can be modular and efficient. The design needs to be optimized to obtain the best results in minimum time and using minimal resources. In the following chapters, these modeling techniques will be used to model performance evaluation system. Because cellular networks are better represented by mathematical modeling and provide more accuracy, MATLAB is used as a designing tool to evaluate the performance of the proposed algorithms. Two systems have been developed to carry out different performance evaluation. The first system allows performance evaluation in terms of handover delay, whereas the second system allows performance evaluation in terms of number of handovers, call dropping, user satisfaction, and minimizing ping-pong effect. The modeling of those systems is also detailed in the associated chapters.

2.4 Summary

In this chapter, a future network and some of the expected entities of the future network are discussed. It is observed that future networks will have entities like CRRM and AAA to better manage heterogeneity of the networks. The chapter also includes a discussion of RRM tasks such as admission control, congestion control, and power control. It is sent that different RRM functions are related with each other to enhance efficiency of the network. Modeling different parts of the network and modeling a performance evaluation system are also discussed. The discussion indicates that, performance analysis system can be designed in many ways to efficiently model different network conditions. These systems should be optimized so that the results can be generated in an efficient and easy way. The optimization of the system can be achieved using abstraction of the functions which do not make any difference to the target of the system.

3 Handover

3.1 *Introduction to Handover*

One of the main objectives of handovers in a cellular network is to provide uninterrupted connectivity between users and the radio access network while they move across cell boundaries within the network. Efficient handover algorithms should be designed to preserve capacity and enhance Quality of Service (QoS) of communication systems in a cost effective manner [1]. There is a number of research works in developing efficient handover algorithms. For example,

3.2 *Causes of Handover*

The occurrence of a handover can be attributed to a number of factors, which could be related to radio link quality, network management and/or service options [4] [6]. While radio link quality related handovers occur frequently and are most difficult to handle, network management and service option related handovers usually occur infrequently and are easy to handle. These causes of handover are discussed in the following sections.

3.2.1 Handover Related to Radio Link

Capacity of wireless communication channel is highly related to the quality of radio link on which the wireless communication takes place. The higher the quality of radio link the higher the capacity of wireless communication can be achieved, and thus better quality of service can be offered. As each user in a cellular network is served on a radio link, it is imperative that the quality of the radio link is maintained to ensure requested Quality of Service to the user. As a user moves away from a Base Station (BS), the radio link quality between the user and the BS becomes weaker. This can be detected by measuring the Received Signal Strength (RSS) at the receiver e.g. base station in uplink communication and/or the Signal to Noise Ratio (SNR) of the received signal. Given that a user is connecting to a BS which is surrounded by a number of BSs, when the RSS of the user goes below a threshold that a reliable wireless connection can be guaranteed or there are other BSs in the vicinity with better RSS available, a handover occurs to pass the user connection from the BS with the weak RSS to the BS with the best RSS, without the service interruption. This type of handover is known as radio link quality related handover [6].

3.2.2 Handover Related to Network Management

As the number of users in a particular cell increases, resource consumption also increases, resulting in congestion within the cell. When congestion in a cell goes above a threshold and the neighboring cells are operating below congestion threshold, users may be handed over to neighboring cells that have lower congestion. This type of handover is known as network management related handover. Apart from this reason, network management related handover can also occur because of a malfunctioning path between BS and a core network through which the user is served. For example, a network may detect some problem in communicating with one BS then it may trigger handover of users in that cell to the nearby cells. Other than this capacity related handover, handover may occur to increase efficiency. For example, if the network detects a shorter path than the current path through which the user is served, then a handover could be triggered.

3.2.3 Handover Related to Service Options

In some cases, a network is designed in such a way that certain services are available only through a particular BS. For example, in the combined network of 2G and 3G, a user can access some data services only when the user is in the coverage of 3G cells. If a user is currently on a 2G network and requests for a data service which is available only via a 3G cell, then the network will handover the user to the nearby 3G cell. Such a handover is called service option related handover [4].

3.3 Types of Handover

Handover can be classified into many types. The classification is presented in the following paragraphs and is also shown in Figure 3.1. Handover occurs between two Points of Attachment (PoA). These PoA can belong to the same or different radio access technologies. When the two PoA are based on the same technology, the handover is known as a homogeneous or intra-system handover, otherwise it is known as a heterogeneous or inter system handover [8].

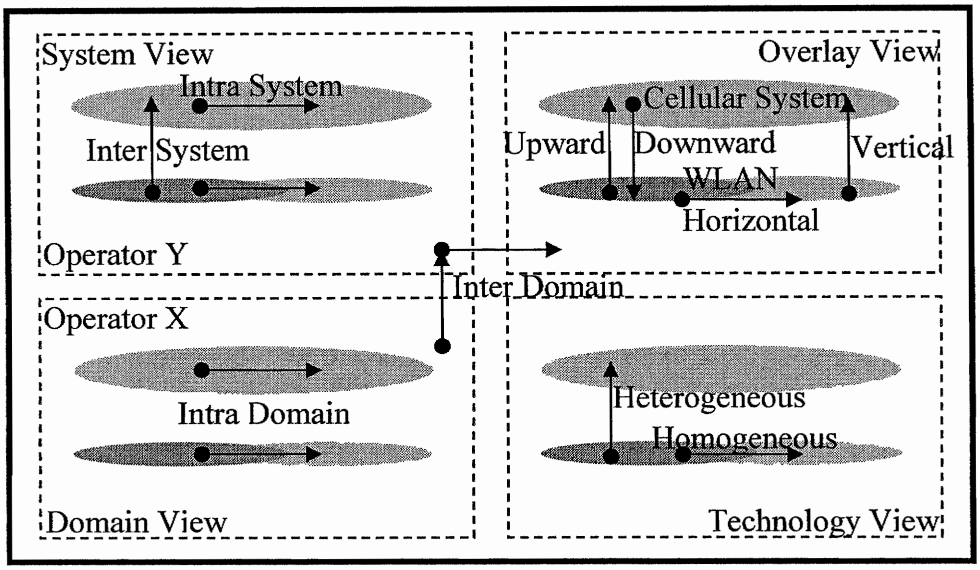


Figure 3.1 Classification of handover [8]

In wireless overlay architectures, it is common to find that more than one network is overlaid above another. In such architecture, when the two PoA belong to the same network overlay, the handover is known as a horizontal handover, and when the two PoA belong to different network overlays, it is known as a vertical handover. If a user is moving from an upper network overlay to a lower one then it is known as a downward vertical handover and the reverse case is known as an upward vertical handover [8].

When an MS can perform a handover between two completely independent domains, which are controlled by different network operators, such a handover is known as inter domain handover, and when the handover is performed within the same domain then it is known as intra domain handover.

In addition to the above classification, which was mainly based on the involvement of PoA, handover can also be classified based on a control point of view. If a handover is controlled by an MS then it is known as a Mobile Controlled Handover (MCHO), and if it is controlled by a network then it is known as a Network Controlled Handover (NCHO). In a case when MS plays a supportive role and a handover procedure is controlled by the network, the handover is known as a Mobile Assisted Handover (MAHO).

Finally, a hard handover is one that occurs when an MS has to terminate a connection before starting a new connection, and a soft handover is one that occurs when an MS can continue with two connections at the same time.

3.4 Desirable Features of Handover

Figure 3.2 summarizes the desirable features that an efficient handover is expected to achieve for improving user experience and service quality.

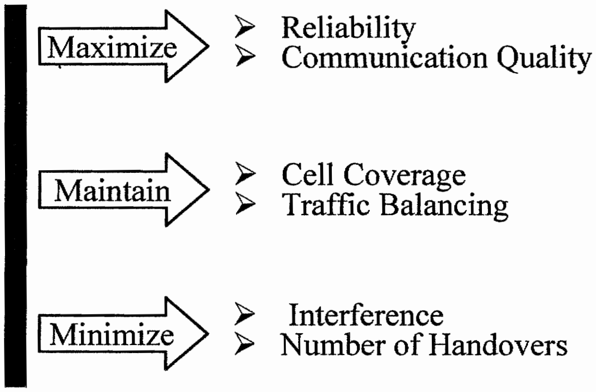


Figure 3.2 Desirable Features of Handover [1]

To maintain a seamless service for a user, handovers should be fast enough so that the user does not experience any disturbance or interruption to the service. The quality of communication link may be degraded due to reducing RSS below an acceptable level or increasing Co-Channel Interference (CCI). For example, a mobile user may move in to a location that the line of sight between the user and the connecting base station (BS1) is blocked, and thus the RSS from BS1 may suddenly decrease. In this case, handing the connection to another base station that can provide higher RSS would maintain the link quality. Radio communication link quality can be maintained by fast handover mechanisms.

One of the most important requirements in cellular networks is that any handover that have been triggered by the networks should be successful. A handover is said to be successful when the call is not dropped while the MS is handed over to the new cell. This can be maximized using efficient channel allocation algorithms and traffic balancing techniques [1].

The effect of handover on the QoS should be kept at minimal. The handover margin should be set so that handover does not result in the QoS decreasing below an acceptance level. If the handover margin is kept high then MS will stay connected with current BS even when signal quality is very low which will decrease QoS.

Congestion within a cell and interference between cells should be minimized, which can be achieved when the handover maintains a planned coverage area of cells. Planning the load of each BS should be efficient. An inefficient planning will likely lead a BS for increasing its power to serve a faraway MS and consequently increasing the interference level in the cell owing to high power.

There should be a check on the total number of handovers. An excessive number of handover increases processing load and signaling which results in the waste of radio resources. Handover in a cellular network is carried out so that users can be served efficiently. However, carrying a handover also require signaling in the network. If there are more number of handovers in the network then it can reduce efficiency of the network by more signaling in the network.

The selection of a target cell should be efficient. Good selection of a target cell leads to decreased number of handovers and increase in communication quality. If a target cell is selected with poor quality, then it is likely that user is far from that BS and there are more chances that user will leave the cell soon. This will in turn increase number of handovers in the network.

Another factor that should be minimized is the signaling involved during handover. This will lead to low handover latency. The signaling in a network takes time, which increases latency. Increase in latency can lead to increase number of call dropping. As discussed above, handover causes signaling in the network which reduces efficiency. Therefore reducing signaling in handover will allow network to serve more traffic.

Handover should have a minimal effect on call blocking. For this, some of the channels can be reserved specially for handover. However, this channel reservation in the cell should be selected carefully as higher number of channel reservation can lead to waste of network capacity and lower number of channel reservation can cause more call dropping in the network. For example, if a network has total N channels. If there are H channels reserved for handover then network can use N-H channels to serve other traffic. If in order to accommodate handover calls efficiently, M channels are required and H > M, then H-M channels are wasted. If H < M then this will cause more call dropping.

3.5 Complexity of Handover

The currently employed handover algorithms perform well only under certain conditions because of the complexities associated with the algorithms. These complexities necessitate better and more efficient handover algorithms. The following sections discuss these complexities.

3.5.1 Structure of the Cellular Network

The performance of a handover is affected in various ways by the different cellular network structures. One of the complex cellular structures is the coexistence of microcells and macrocells. Microcells serve hotspot areas, whereas macrocells cover low density areas. To obtain good performance, different cell structures need different handover algorithms [50]. In a wireless overlaid architecture, a cellular structure is overlaid with both microcells and macrocells. In such an overlaid architecture, microcells are served with high bandwidth, low coverage technologies such as WLAN, whereas macrocells are overlaid on this cell structure supported by relatively low bandwidth, high coverage technologies such as 2G, 3G, or 4G. As the cell size decreases, the number of handovers per average call duration increases, while the variables such as RSS, SIR, and Bit Error Rate (BER) will change faster and the time available for performing a handover decreases [51].

3.5.2 Topographical Features

A signal is characterized by the magnitude of propagation loss exponent and breakpoint. (Breakpoint is defined as a point at some distance where the exponent related with path loss changes.) The performance of a handover depends on the signal profile in the region served by a network [53]. Topographical features of the coverage area which changes signal profile adds more complexities to the design of a handover scheme. If the signal profile in a region is changing rapidly then it becomes more difficult to make a handover decision. Handover design will also need to accommodate change in signal profile which will make handover procedure more complex.

3.5.3 Traffic

Traffic distribution is a function of location and time. The network employs various approaches to balance traffic among adjacent cells to cope with traffic non uniformities. Some of the approaches use different cell size, non uniform channel allocation in different cells, and dynamic channel allocation [53]. In order to achieve a balanced traffic across all cells, different handover policies are incorporated.

3.5.4 Radio Propagation Phenomena

Radio propagation highly depends on the surrounding environment. In certain complex environments, radio propagation characteristics can vary rapidly, e.g., in a city with high buildings close to the sea, and therefore it needs a smart handover algorithm to cope with such uncertain situations. Other environmental characteristics such as street corner effect also add complexities to a handover [6]. The environment dependent handover algorithm performs with higher efficiency in these cases [8].

3.5.5 System Constraints

Every cellular system has its own power control algorithms, which can significantly affect the characteristics of the handover algorithm. For instance, some systems allow the user to stay connected using minimum required power transmission whereas others handover the call to achieve higher performance [11]. This interaction between power control and handover also adds complexities to handover scheme.

3.5.6 Mobility

Networks are designed depending on the different levels of mobility. Networks such as WLAN are designed for low mobility users, whereas cellular networks such as GSM, CDMA, and UMTS are designed for a wide range of users, ranging from stationary users to high speed users (up to speeds such as 120kmph) [6]. For high speed users, a handover has to be performed much quicker to ensure uninterrupted sessions.

As new networks evolved, the complexities in a network increase. To cope up with these increasing complexities, handover procedures must improve as well. In Next Generation

Networks (NGNs), handover algorithms has to be smart enough to understand the user and network requirements to perform effective and seamless handover [1].

3.6 Handover Parameters

In this section, different parameters involved in making a handover decision are described. The conditions of a network can be understood by studying the different network and user parameters.

3.6.1 Received Signal Strength

RSS is an important and useful parameter that is widely used in handover decision making. A close relationship exists between RSS and the distance between BS and MS. RSS is affected by topological features such as buildings and it ignores the effect of CCI. The coverage area would become smaller than the planned one because of the topological obstacles. Using only the RSS value as a handover criterion can lead to excessive number of handovers because it can trigger ping pong effect [1].

3.6.2 Signal to Interference Ratio

SIR is a common parameter associated with voice signal quality, system capacity, and dropped call rates. It gives a clear picture of the current quality of a network with respect to the receiver i.e. MS in downlink communication. SIR is estimated using BER. When actual SIR is lower than designed SIR the handover has to be performed fast which is not always possible. Therefore, this will increase a number of dropped calls in the network. SIR also determines the reuse distance. Reuse distance is the distance at which same frequency band can be reused. At the point where SIR decreases to the level that it would not create interference, the same frequency can be used again. Some of the disadvantages of SIR are that (i) it can oscillate as a result of the propagation condition and can cause ping pong effect and (ii) although BER is a good indicator of link quality and useful to be used as a handover parameter. However, it can be high exactly below the BS as well, in which case the handover may not be suitable [54]. Limited interference deterioration in BER will not necessarily require an inter cell handover. An intra cell handover may be sufficient in this case [52].

3.6.3 Transmission Power and Other Parameters

In uplink communications where MS transmits signal to the BS, the transmit power of MS needs to be considered carefully as it is directly related to MS battery life. Choosing a base station to where a MS can communicate using low transmission power plays an important role in reducing power consumption. This criterion triggers the handover to a BS which calls for minimum uplink transmission power. Some of the other parameters are traffic, call and handover statistics, and velocity. These parameters allow us to make a handover decision to reduce the total number of handovers and manage the network more effectively [1].

3.7 Conventional Handover Algorithms

Most of the conventional handover algorithms deal with homogeneous networks only and are mainly based on parameters such as signal strength, SIR, distance, velocity, direction, power and BER. For an effective handover, different approaches have been proposed. These approaches have their own advantages and disadvantages and each of them performs better than the others under certain circumstances. These approaches are described in the following sections.

3.7.1 Signal Strength Based Handover Algorithm

Signal strength based algorithms are classified into signal strength based, relative signal strength based and absolute and relative signal strength based algorithms.

According to the RSS criterion [14], an MS is connected with a BS from which it receives the highest signal strength. This algorithm allows the MS to be attached with the BS with the strongest signal. The disadvantage of this algorithm is that the signal strength may vary because of shadowing and propagation characteristics, e.g., corner effect, which can cause frequent handovers. Another disadvantage is that the MS may continue to stay connected to the current BS until it receives good signal strength from other base stations even if it has crossed the designed cell boundaries. This will lead to difficulties in maintaining physically planned cell boundaries.

The addition of hysteresis as an additional criterion will help to overcome these disadvantages. Although hysteresis helps in preventing frequent unnecessary handovers, it does not help in reducing call dropping as it introduces delay in necessary handover [58]. A balance between the number of handovers and delay in handover should be achieved by averaging signal strength and taking the appropriate hysteresis. By introducing hysteresis, if the RSS of any other BS exceeds the RSS of the current BS, then a handover is performed to the new BS. This can cause ping pong effect [58]. To avoid this, a handover margin is set. The North American Personal Communication Systems (PACS) and Personal Communication Service (PCS) standards combine hysteresis with dwell timer [49] to decrease total number of handovers.

In the absolute signal strength algorithm, a handover is requested when the RSS drops below the threshold of -100dBm for a noise limited system and 95dBm for an interference limited system [6]. The handover threshold can be varied to achieve better performance. The threshold value is set by observing the slope L of the RSS curve against distance and the Level Crossing Rate (LCR) of RSS. If L or LCR is high, it implies that the MS is moving away faster, in which case the handover has to be performed faster. If L or LCR is low, it means that the MS is moving slowly and hence the handover initiation threshold can be kept accordingly. The handover threshold can also be determined dynamically by the mobile velocity and path loss slope. This will help completing a handover successfully and avoid unnecessary handover. However, this algorithm has several serious disadvantages. BS increases transmission power to cope with high interference. If the RSS is very high because of high interference, then a handover will not take place, although ideally, a handover is desirable to avoid interference. If the RSS is low, a handover will take place even if the voice quality is good, although ideally, a handover is not required. In such cases, the Supervisory Audio Tone (SAT) information is used along with the RSS by some systems to avoid unnecessary handover [1].

If the absolute and relative signal strength handover algorithms are combined, a handover will take place when the following two conditions are satisfied [59]: the average signal strength of the serving BS should be below an absolute threshold T dB and the average signal strength of candidate BS should exceed the average signal strength of the current BS by an amount of h (hysteresis) dB. The first condition prevents the occurrence of a handover when the current BS can provide sufficient signal quality. Beck [60] has shown that an optimum threshold achieves the narrowed handover area (and hence reduces interference) and a low expected number of handover. The basic variables for this handover algorithm are the threshold level and the hysteresis margin [5]. Some of the findings reported in reference [50] with respect to this algorithm are (i) the probability of not finding a handover candidate channel decreases as the overlap region increases, (ii) the probability of not finding a handover candidate increases as the handover threshold increases, (iii) the probability of a late handover (handover occurred after the optimum time of handover) decreases as the handover threshold increases, (iv) the probability of unnecessary handover, i.e., the ping pong effect, increases as the handover threshold increases and (v) the probability of unnecessary handover decreases as the hysteresis increases.

3.7.2 SIR Based Handover Algorithm

SIR is one of the measures of communication quality. To ensure good quality voice, SIR at the cell boundaries should be targeted at relatively high, e.g., 18 dB for Advanced Mobile Phone System (AMPS) and 12 dB for GSM. However, a lower SIR target level may be used for capacity reasons. If a lower SIR is used, the number of cells per cluster is small and channels can be reused more frequently in a given geographical region [8]. This algorithm performs a handover when the SIR of the current BS drops below the threshold and at the same time another BS can provide sufficient SIR. Hysteresis can be incorporated in this algorithm also. In either case, a handover is desirable. However, SIR based handover algorithms prevent a handover near nominal cell boundaries and cause cell dragging and high transmission power [61]. Chuah [61] has suggested an uplink SIR based algorithm for a power controlled system. In this system, each user tries to achieve a target SIR and a handover occurs when the user's SIR drops below a threshold, which is normally less than the target SIR.

3.7.3 Velocity Based Adaptive Handover Algorithms

Velocity based adaptive handover algorithm helps in handling handover for users with different speeds. When users are moving at different speed, their handovers should also be performed in different time. A handover request from fast moving vehicles must be processed quickly, and this can be done using a handover algorithm with short temporal averaging window. However, if the length of the average window is kept constant, optimal handover performance will be obtained only at a particular speed. Velocity adaptive handover algorithm provides a good performance for an MS with different velocities by adjusting the effective length of the averaging window [58]. This algorithm can also serve as an alternative to the umbrella cell [58] approach for tackling high speed users if low network delay can be achieved. As the umbrella cell requires extra infrastructure, this alternative approach can lead to savings in terms of infrastructure cost.

One of the velocity estimation techniques uses the Level Crossing Rate (LCR) of the RSS in which the threshold level should be set as the average value of the Rayleigh distribution of the RSS [59], requiring special equipment to detect the propagation dependent average receiver power. Kawabata [59] proposes a method of velocity proportionality to the Doppler frequency. The velocity estimation technique exploits the method of diversity reception. If an MS is using selection diversity, this method requires no special equipment.

The characteristics of a velocity adaptive handover algorithm for microcellular systems are presented in reference [59]. Three methods for velocity estimation are analyzed: the LCR method, zero crossing rate method, and covariance approximation method. It is found that the spatial averaging distance that is required to sufficiently reduce the effects of fading depends on the environment. This algorithm can adapt the temporal averaging window (a window with a certain time length) used to take samples of RSS value. The window can be adapted either by keeping the sampling period of LCR constant and adjusting the number of samples per window or vice versa.

3.7.4 Direction Based Handover Algorithm

If an MS moves fast and is not handed over quickly enough to another BS, the call may be dropped. In a Non Line Of Sight (NLOS) handover, MS may experience an undesirable condition called the corner effect [62]. If a user is moving across the corner then that will cause abrupt decrease in the signal which is known as corner effect. This can be solved by either connecting the fast moving vehicles to an umbrella cell or using better handover algorithms such as a direction based handover algorithm [62]. A direction based handover algorithm improves cell membership properties and the performance of handover scheme in Line Of Sight (LOS) and NLOS scenarios in a multi cell environment. A handover algorithm is said to possess a good cell residency time if the probability that an MS is assigned to the closest BS is high. The term "closest BS" refers to the BS which is closest to the MS throughout the call duration [62]. Improvement in cell residency time leads to fewer handovers and reduced interference.

The direction based handover algorithm works on the basic principle that a handover to the BS toward which the MS is moving is encouraged, whereas a handover to the BS from which the MS is receding is discouraged. The advantages of using this algorithm are that the probability of dropped calls for hard handovers is reduced, e.g., in TDMA systems, and the time needed by a user to be connected to more than one BS for soft handovers is also reduced, e.g., in CDMA systems, allowing more potential users per cell.

A variation of the basic direction based algorithm is the preselection direction based algorithm [62]. In brief, if a BS with the best RSS is a receding one and has quality only slightly better than the BS with the second best RSS and the BS with second best RSS is approaching, the handover should be made to the second best BS because it is more likely to improve its chances of being selected [62]. This provides a fast handover algorithm with good cell membership properties without the undesirable effects associated with large hysteresis.

3.7.5 Minimum Power Based Handover Algorithm

A minimum power based handover algorithm, which minimizes the uplink transmission power by searching for a suitable combination of a BS and a channel, is suggested in reference [61]. Although this algorithm reduces call dropping, it increases the number of unnecessary handovers. This high number of handovers can be avoided by using a timer such as dwell timer. First, the channel that gives minimum interference at each BS is identified. Second, the BS that has a minimum power channel is determined. Mende [63] uses a power budget criterion to ensure that the MS is always assigned to the cell with the lowest path loss, even if the thresholds for signal strength are not reached. The advantages of this criterion are low transmission power required and a reduced probability of high CCI.

3.7.6 RSS and BER based Handover Algorithm

An algorithm based on RSS and BER is described in reference [64]. With regard to RSS, a threshold is used for the current BS, and a hysteresis is used for the target BS. With regard to BER, a separate threshold is defined. The target BS can be either included or excluded from the handover decision process. The later scheme is employed in GSM in which the MS does not know the signal quality of the target BS. In principle, it is possible to measure BER of the control channel of the target BS. The three parameters considered in the simulations in Ref. [64] are RSS threshold and RSS hysteresis window size. In general, a low BER threshold value will reduce the probability of handover requests. The most efficient BER value that can have optimal number of handovers is when BER from both BS are equal. However in real scenario, the value of BER changes from one BS to another. In this case, if BER value from current BS is used, then handover will occur before the optimal time. If the BER value of target BS is used that that leads to handover after the optimal time. In this case, hysteresis can help in performing the handover at correct time.

3.8 Emerging Handover Algorithms

As a consequence of the growing network complexity and service requirements, conventional handover algorithms are not able to meet new requirements such as.... The evolution of new networks has also triggered the evolution of novel handover algorithms. These emerging handover algorithms are discussed in the following sections.

3.8.1 Dynamic Programming Based Handover Algorithm

A dynamic programming based handover algorithm provides systematic solution to the handover problem. However, the efficiency of the handover algorithm depends on the model used. This algorithm is described in detail in reference [65]. The algorithm uses various recursive routines to compute optimal solutions when signal strength drops below the required level. Handover as a trade off between service failure events and number of handovers has been described in reference [66]. The solution to this problem has been derived using the dynamic programming and recursive algorithms, which call for complex computation and tracking mobile trajectory to make effective decisions. A handover framework has been proposed as a stochastic control framework in reference [67]. The algorithm described in reference [67] models RSS samples of the MS as a stochastic process. The handover is described as reward/cost process where the cost represents the resources needed to perform handover and the reward refers to a function based on various characteristics such as signal strength, channel fading, shadowing, power control strategies, traffic distribution, cell congestion profile, and channel assignment. The handover has been described as a switching penalty based on needed resources to perform a successful handover. Simulation results presented in reference [65] show that this algorithm achieves better performance than the relative signal strength based handover algorithm.

3.8.2 Pattern Recognition Based Handover Algorithm

This algorithm is discussed in detail in reference [68]. It provides a solution to handover by predicting the handover of the mobile node. Pattern recognition allows deriving meaningful regularities in a complex environment. Two techniques are described for pattern recognition: explicit and implicit. The former technique applies the discriminate functions that define (n-1) hyper surfaces in an n-dimensional feature space, whereas the latter technique measures the distance of current pattern with predefined patterns from representative class. This distance allows the prediction of the moving pattern of current user and allows making the handover decision according to the pattern of the user. Explicit pattern recognition technique defines (n-1) hyper surfaces, and input pattern is classified according to locations of MS on these hyper surfaces. To minimize the uncertainty of the algorithm, efficient processing of RSS measurement is needed. An inefficient processing of RSS measurement can result in unnecessary handovers.

3.8.3 Prediction Based Handover Algorithm

Prediction based handover algorithm is proposed in reference [69]. This algorithm predicts the value of handover related criteria such as RSS and SIR. According to [69] the technique performs better than the relative signal strength based algorithm and the combined relative and absolute signal strength based algorithm. An adaptive prediction based algorithm is used to achieve trade off between the number of handovers and signal quality [69]. This algorithm can use path loss and shadowing as prediction factors. Path loss depends on the distance, whereas shadowing is based on correlation between distance and property of surrounding. The value of both path loss and shadowing can be calculated using correlation analysis, which is used to predict the value of different criteria related to handover. The future value of RSS is calculated based on the previously measured RSS using adaptive Finite Impulse Response (FIR) filter, which can minimize unnecessary handovers.

3.8.4 Fuzzy Handover Algorithm

Fuzzy logic based algorithm is proposed in Reference [1]. This algorithm uses fuzzy logic to make handover decision. The complexities involved in this handover criterion are tackled using fuzzy logic. The idea is to take handover decision using fuzzy value of the parameters. This algorithm is explained in detail in Chapter 5. The algorithm performs better than the traditional signal strength based algorithm. It weighs different parameters according to the needs of the network. The algorithm is also capable of handling heavy fading. Fuzzy logic has inherent fuzziness capable of handling overlapping of the cells and criteria. The solution allows to measure then membership of the cell for handover. This approach can reduce the number of handovers.

This section presented some of the emerging handover algorithms. It is interesting to note that as the complexity of networks increases, handover algorithms are also needed to be smarter. In future a handover will be more complex and based on more criteria because of the involvement of more than one network. Such algorithms are discussed in detail in Chapters 4 and 5. Handover, being one of the most important radio resource management tasks, has a close relationship with other RRM tasks. This engagement is also modified to achieve better performance of the network in addition to techniques such as handover prioritization, interworking with power control, channel allocation, and other RRM functions. In the following section, the study of other RRM functions and their correlation with handover is discussed.

3.9 Handover Prioritization

The main use of handover prioritization is to reduce the total number of handovers. Handover prioritization can be used with various other RRM tasks such as admission control, power control, and channel allocation. The two main methods of handover prioritization are guard channels and queuing.

3.9.1 Guard Channels for Prioritizing Handover

The guard channel approach of prioritizing handover reserves few channels exclusively for handovers. This approach allows prioritizing handover over other calls, thereby reducing the call dropping rate. If N channels are reserved for handover from a total of C channels in the cell then the remaining C–N channels are shared between handover calls and routine calls [1]. Handover fails only when no channel is available in the cell in the reserved quota of handover and normal channels. Dropping probability can be reduced by this approach. For an efficient prioritization scheme, the optimal number of guard channels in the cell must be reserved. An excessive number of guard channels can reduce the use of channels and can result in higher call blocking rate in the cell.

3.9.2 Queuing of Handover

Queuing is the approach of delaying handover instead of dropping the call when handover fails [1]. In this approach, the call is continued albeit with slightly reduced quality rather than dropping the call completely. When a handover attempt is about to fail because the channel is unavailable, instead of dropping the call, the call will be continued in the current cell and handover will be executed after some time. The timer for holding a handover call with low quality can be adaptive or an exponential timer which can have higher limit according to the path loss slop of the cell can be used. The signal quality will be slightly reduced during the delay of handover but the call dropping rate will be reduced with this approach.

The queuing approach is not necessary when handovers occur uniformly. This approach will be useful when there are handover requests in group and serving all the requests simultaneously is not possible. Queuing can be effective when the traffic is low the handover requests are more. If the traffic is high, then it is very unlikely that a queued handover request will be entertained.

3.10 Summary

This chapter provides a study of handovers. The contents include an overview of handover, its classification, the complexities involved, the desired features of handover, the parameters used to make a handover decision and the handover algorithms. It is identified that the handover algorithms are becoming more complex, and in heterogeneous networks they become even more complex. It is also shown that different techniques to give priority to handover or other RRM tasks are used to improve allover performance of the network.

4 Network Controlled Vertical Handover

4.1 *Introduction*

As discussed in Chapter 1, 4G systems will be a convergence of different technologies, with IP as the common layer between the different technologies. As a result, most of the schemes for a vertical handover specified in previously published literature are based on MIPv6, which has been designed to support the interoperability between such technologies. In traditional wireless systems, the signaling involved in any handover scheme is critically important for the overall delay in handover. In 4G systems, the issue becomes much more challenging because of the nature of the interoperability between different technologies. In this chapter, MIPv6-based handover scheme, Mobile Controlled Vertical Handover scheme (MCVHO) and the newly proposed Network Controlled Vertical Handover Scheme (NCVHO) are described. An analysis of the performance in terms of handover delay and signaling in all three schemes is also presented.

4.2 *MIPv6 Based Handover*

MIPv6-based vertical handover process consists of three phases: (i) discovery or scanning phase, (ii) re-authentication phase, and (iii) re-association phase which are described in the following sections. In a discovery phase, candidate PoAs are discovered by scanning a number of channels. In a re-authentication phase, the MS is authenticated with the target PoA. Finally, the MS is associated with target PoA so that it can start connection via that PoA. Figure 4.1 presents the signaling of MIPv6-based vertical handover in accordance with the specifications in Reference [25].

In Figure 4.1, MIPv6 based handover is shown. The first phase, namely discovery phase, is performed using router solicitation and router advertisements. This part of the procedure results in network evaluation and selection of the target Point of Attachment (PoA) to perform handover. This phase is followed by a re-authentication phase. In the re-authentication phase, the authentication of an MS with a new PoA (nPoA) is performed using binding update and binding acknowledgment messages. After the completion of this phase, association with the nPoA is performed. Because of the heterogeneity of the networks, Handover (HO) and Care of Address (COA) tests are included in MIPv6-based vertical handover. Finally, the connection is transferred to the nPoA and the connection with the old PoA (oPoA) is closed. In the following sections, the three phases of an MIPv6-based vertical handover are described.

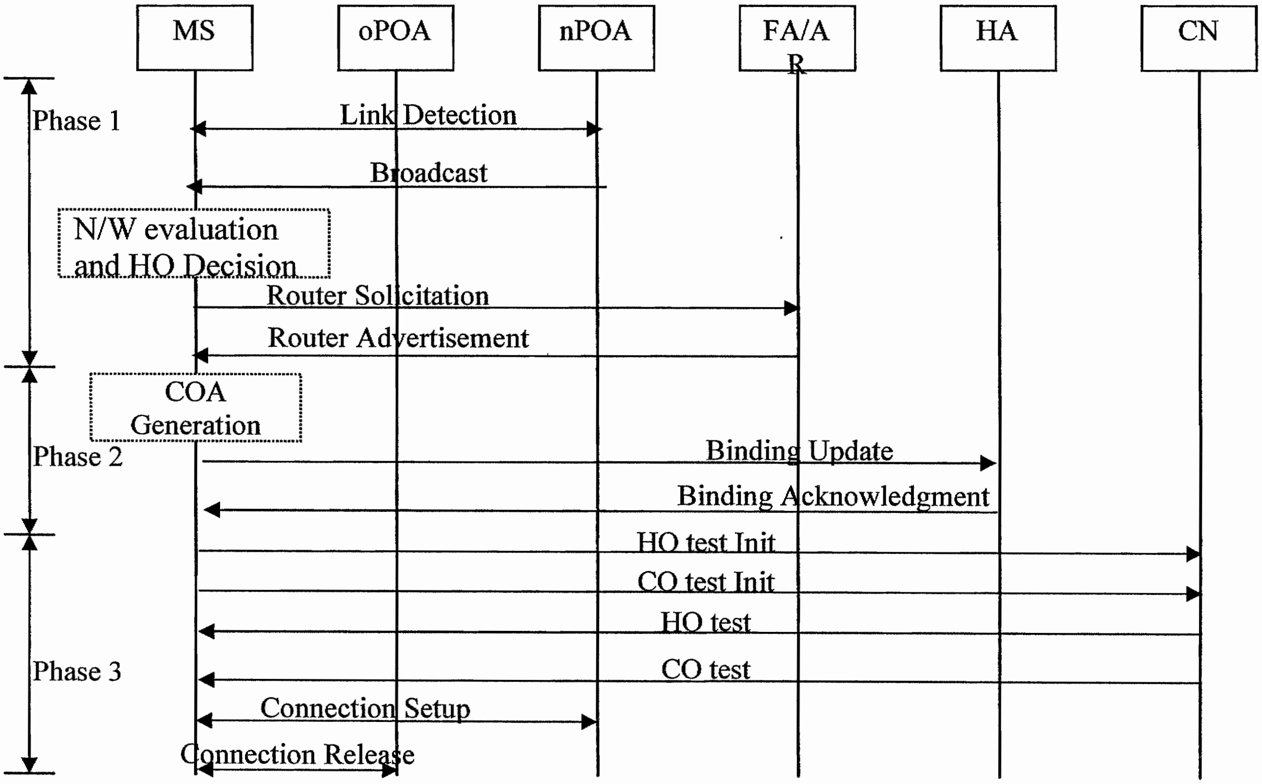


Figure 4.1: Handover using MIPv6 [25]

4.2.1 Discovery or Scanning Phase

To trigger the discovery of an nPoA, an MS needs to have certain criteria for initiating the handover process. Usually, when the strength of the signal received by the MS goes below a signal quality threshold, it starts scanning for all the PoA in the vicinity. In order to perform this scanning, the MS can perform either an active or a passive search. In an active search, MS sends probe requests on different channels and in a passive search it just listens on different channel for the beacons. It also decides whether to perform a complete search or a partial search. A partial search can be carried out to reduce the timing. A channel list is generated for scanning. In a complete search, probe requests are sent on all the channels are included in a list and in a partial search, only randomly selected channels are included in a list. The MS sends requests on the channels in the list, to get an indication that the PoA is willing to accept the request. The MS creates a report of all PoA who have sent a probe response.

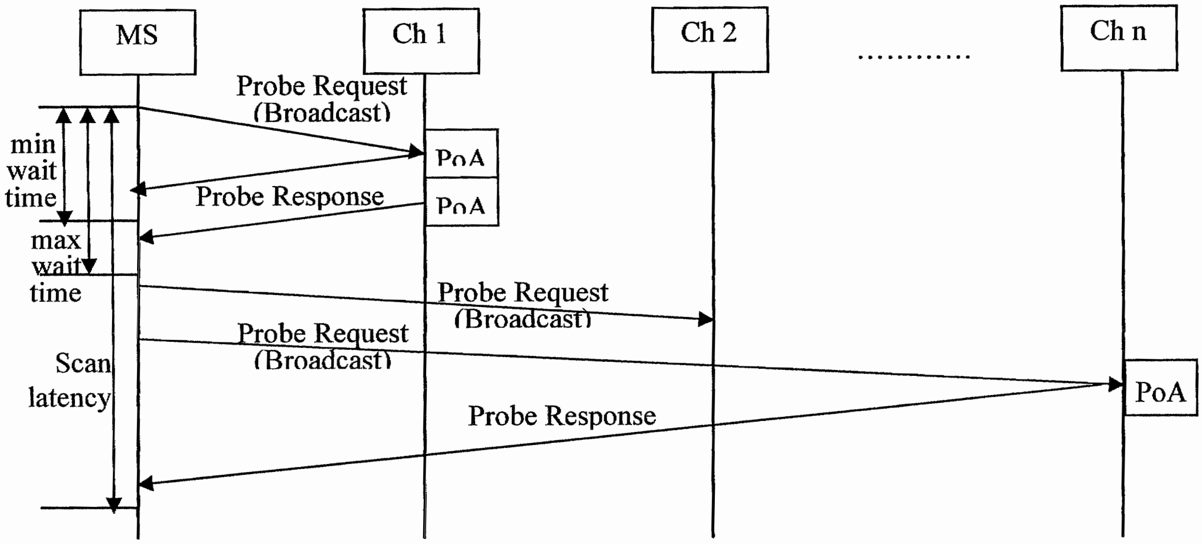


Figure 4.2: Discovery or scanning phase.

The parameters that control the discovery phase are *ScanType, ChannelList, ProbeDelay, MinWaitTime, and MaxWaitTime.* This discovery procedure is show in in Figure 4.2 and described in the following algorithm from Reference [25]. In the first step, it scans each channel in the channel list to detect activity if the activity is not detected the channel is skipped. Once, it completes detecting each channel then it broadcasts probe request on each active channel form the channel list. For the first time, it sets timer to *minWaitTime.* If there is no activity during this time period then it scans the channels with timer value set to maxWaitTime. As a result of this process a list of PoA is generated from which MS has received probe responses.

Procedure: Full or short scan in active discovery.

```
For each channel in channel_list do,
     Tune MS in channel to probe.
     wait for channel activity detection or
     Probe_delay timer expiration.
     If ProbeDelay_expires, then
          Probe_next_channel
     If channel_activity_detected then
          MS gets access to wireless medium.
MS broadcast probe_request_frame.
MS wait MinChannelTime sensing the channel.
If no_activity_detected then
     Probe_next_channel
If traffic_detected or probe_response then
     wait Max_Channel_Time sense channel.
do process_received_probe_response.
Until MaxChannelTime expires, then
     Probe_next_channel
```

4.2.2 Re-Authentication Phase

After the MS has its list of candidate PoA, then the MS has to get authenticated with the selected nPoA. This phase is the prerequisite of re-association phase. To perform re-authentication, the MS sends an authentication request, which includes credentials and other service details from the oPoA such as secret key. In case of MIPv6, this authentication is carried out via a Home Agent (HA) of the MS. Credentials provided by the MS are sent to the HA by the nPoA for verification. This procedure is shown in Figure 4.3. MS sends a binding update message via nPoA to HA. HA replies with binding acknowledgement message to confirm authentication.

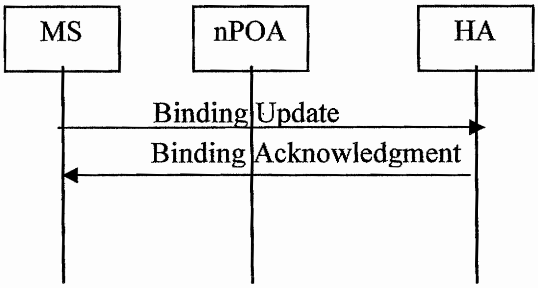


Figure 4.3 Re-authentication Phase

Although authentication is a prerequisite of the re-association phase, the exact point of time for authentication is not defined by the standard. Some vendors take advantage of this fact to reduce handover latency. Techniques such as pre-authentication are used for this purpose. Examples of such techniques are discovery with pre-authentication, in which the authentication of an MS is performed as soon as the discovery phase is completed. IAPP-based pre-authentication is another example of re-authentication, in which an MS is authenticated with all PoA in the Extended Service Set (ESS) as soon as the MS is associated with the first PoA in that ESS. This is performed well before the discovery phase has started. In this case, re-authentication does not contribute to handover latency. The authentication information is transferred using the Distribution System (DS) to all PoA in the ESS so that re-association can directly follow the discovery phase.

4.2.3 Re-Association Phase

After the discovery of candidate PoA and authentication with nPoA are accomplished, the MS re-associates with the nPoA. This phase is shown in Figure 4.4 and the handover is completed in this step.

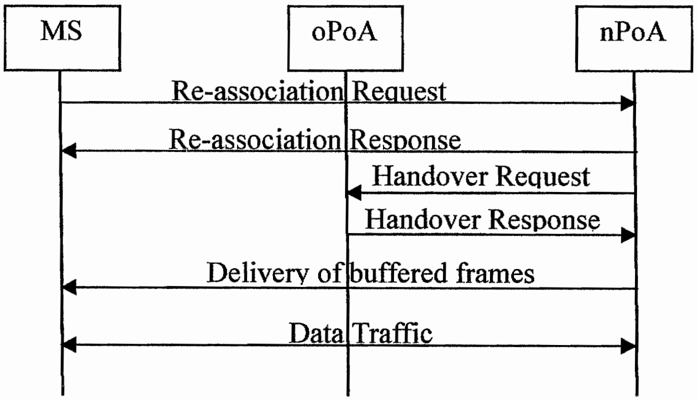


Figure 4.4: Re-association phase

The MS sends a re-association request to the nPoA which in turn sends a re-association request to the oPoA. The oPoA sends a re-association response to the nPoA which complete the association between oPoA and nPoA and the nPoA initiates HO and COA tests. These tests are necessary when the oPoA and nPoA belong to different networks. If the nPoA and oPoA are in the same network, these tests are skipped. The HA is notified about the completion of the handover after which the HA forwards all buffered packets to the nPoA and the connection with the oPoA is closed. What happens to the packets buffered in oPoA

As discussed earlier, MIPv6-based handover has evolved to meet network evolution, and some signaling is added to deal with heterogeneity of the networks involved in handover. However, some of the issues involved in an MIPv6-based handover are still creating bottlenecks in improving handover performance. For example, in MIPv6-based handover, authentication is always performed via HA which adds significant delay in the procedure. Such issues are discussed in detail in the following part of this chapter. Before that, other vertical handover schemes are described next.

4.3 Mobile-Controlled Vertical Handover (MCVHO)

The MCVHO scheme is proposed in Reference [7]. Although the MCVHO scheme uses most of MIPv6 signaling to perform vertical handover, it enhances basic MIPv6 handover. Reference [7] proposes a vertical handover between UMTS and WLAN in two cases. First, when networks belong to the same administrative domain; second, when networks belong to different administrative domains. Signaling in MCVHO is shown in Figure 4.5.

In a situation when an MS decides to perform a handover from one network to another, it sends a vertical handover request to the target PoA in the new network. The request consists of the IP address of the MS. The target PoA checks if the IP address belongs to the its administrative domain; then the PoA will send a vertical handover reply to the MS indicating whether it is in the same domain.

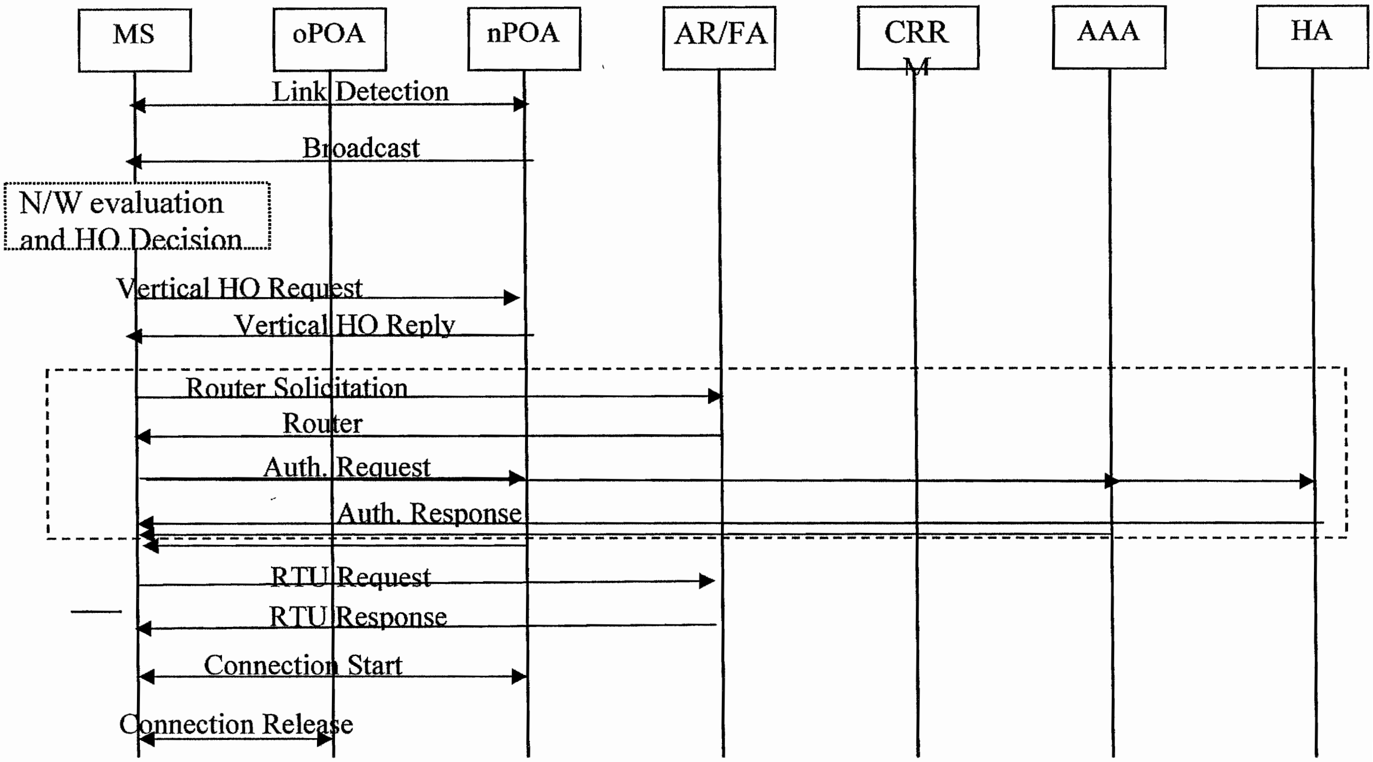


Figure 4.5 Mobile Controlled Vertical Handover

After receiving the vertical handover reply, the MS checks for the type of vertical handover it has to perform. If the target network belongs to the same administrative domain, then the MS can either directly send a routing update (RTU) request in the First case. In second case (i.e., when networks belong to different administrative domains), MS has to perform an MIPv6-based binding update to get its new care-of-address or home address. This extra signaling of vertical handover is shown in Figure 4.4. with a dotted box

In this case, once the MS performs MIPv6 signaling it gets a new IP address through the vertical handover reply message from the nPoA. The MS now sends a routing table update message to the router; this message includes the current IP address of the MS, the IP address of the nPoA, and network type of the target network.

On receiving the RTU request by the router, it knows the new path of the MS along with network type so that it can correctly format packets. Once the router agrees for the handover according to the RTU request, it sends an RTU reply to the MS as an acknowledgement of the RTU request.

The MCVHO proposed in Reference [7] depicts a vertical handover scheme by enhancing MIPv6 with direct authentication in case of a intra-domain handover; in other case, it has to follow the MIPv6 procedure for binding update.

4.3.1 Summary

MIPv6-based vertical handover and MCVHO are described in the earlier section. These two methods use mobile controlled handover. However in a mobile controlled handover, the MS has limited battery power and it also has very limited information about current network conditions such as network traffic information, network capacity and current Radio Resource usage information. These limitations of mobile controlled handover are highlighted in Section 4.5. In the above-specified handover procedures, there is a large amount of signaling on the wireless part of a network. This signaling increases handover latency considerably. If this signaling can be reduced, then it can reduce handover latency. This approach is taken for designing the new handover scheme presented in the following text.

4.4 Network-Controlled Vertical Handover (NCVHO)

Network Controlled vertical handover allows network to perform a handover on behalf of MS. This technique uses more information from CRRM to enhance the network selection capacity.

In the NCVHO scheme, the control of the handover procedure is handed over to the PoA. The MS plays a supportive role in performing the handover. The complete procedure is shown in Figure 4.6. The scheme is based on an original MIPv6-based handover but now MS is not involved in most of the handover signaling. This scheme is described in the following paragraphs.

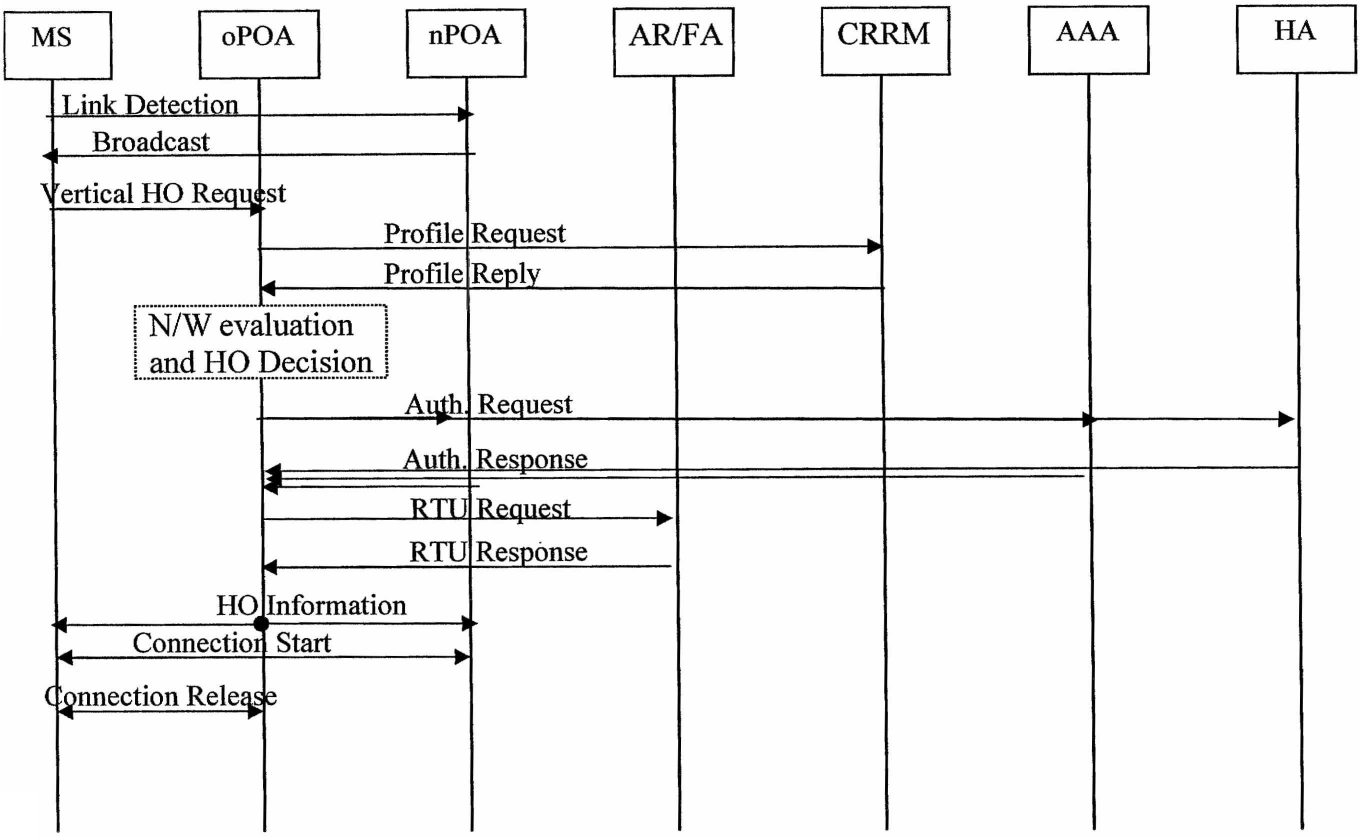


Figure 4.6 Network-Controlled Vertical Handover

When the MS is connected with a PoA, it detects beacons from other PoA in the vicinity. When the MS wants to handover to an nPoA, it sends a vertical handover request to the oPoA. This request contains a list of PoAs in order of preferred connectivity based on user preference as well as information such as signal strength and IP addresses of candidate PoA.

Upon receiving a vertical handover request from the MS, the oPoA sends a profile request to the corresponding CRRM entity. The CRRM entity, as discussed in Chapter 3, has network performance-related information of different networks in the area. This information or profile may include network load, available bandwidth, and network-related parameters which are not known to either the MS or the individual RRM entities. After assembling this information, the CRRM entity sends these profiles as a profile reply to oPoA.

After the oPoA gets a profile response from the CRRM entity and information from MS, it can now evaluate the list of candidate networks. This evaluation process can be carried out using any of the decision procedures proposed in References [28, 29, 1, 38], taking into account network parameters and user preference.

After the nPoA is selected to make a handover, authentication of the MS with the nPoA has to be performed in one of three ways. First, if both the oPoA and nPoA belong to same domain and a direct association exists between nPoA and oPoA, then the authentication of the MS with the nPoA can be performed directly by the oPoA. Second, when the oPoA and nPoA are in different domains but are connected via a common AAA architecture, then authentication can be performed via an AAA agent between oPoA and nPoA. In a third case, no association exists between the oPoA and nPoA (i.e., the nPoA does not trust the information provided by the oPoA). In this case, as the MS is able to access services from nPoA, both nPoA and oPoA must have some kind of association with the home network of the MS; authentication of the MS with nPoA in this case needs to be done using conventional MIPv6 binding update procedure.

After the oPoA completes the authentication of MS with nPoA, oPoA receives binding information of MS.. The oPoA sends an RTU request message to the corresponding router of MS to update the routing information of MS. This RTU request includes the new IP address of MS and the IP address of nPoA.

The router will check if the new network address of the MS complies with the new network (nPoA). After the RTU approves this new route of MS, it updates the routing table, and the packets are then routed to the MS via the nPoA at the completion of handover. The router notifies the oPoA about this new route of MS by sending an RTU reply message.

Now the oPoA has binding information of MS, and in addition to that it has also updated the routing table of the router. The oPoA can now send a handover command to the MS and the nPoA. This handover command includes binding information and the new IP address of MS which allows the MS to start a connection directly via nPoA. Once a connection with the nPoA is established, the connection with the oPoA is closed.

4.5 Performance Analysis

All the three handover schemes described in this chapter try to improve the performance of the handover procedure. The performances of these schemes under various situations differ because they have alternative approaches of performing handover. In this section, the performance of these handover schemes are analyzed in terms of amount of signaling, usage of radio network, and handover delay; the performance with respect to other parameters such as call dropping, number of handovers, and user satisfaction are also analyzed.

The performance can be analyzed using analytical method, simulation, or emulation. For analyzing the performance in terms of amount of signaling, the analytical method, which is based on signaling presented in the previous section, is used. Performance analysis of all three methods in terms of handover delay is done using simulation. The analysis is presented in the following section.

4.6 Signaling Delay in Handover

To compare the delays in the performances of the different handover schemes, the method specified in Reference [30] is used. In the first part of this analysis, a simple model shown in Figure 4.4 is used. In this model, each link between the nodes in the network has certain delay. This delay specifies the time that a signal will take to traverse from one node to another. The total delay of a signal traversing from Node X to node Y will be the sum of all link delays that form a path from Node X to Node Y. The total delay of a signal increases when the number of nodes in the path increases. The delay associated with an individual link is shown in Figure 4.7.

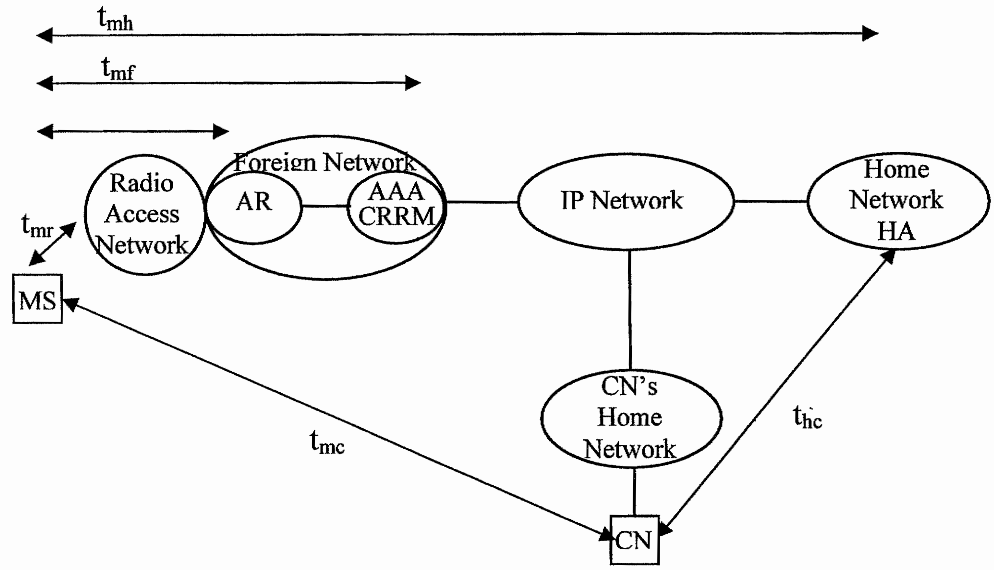


Figure 4.7 Link delay model [30]

The delay notations used in Figure 4.7 are described as follows.

- The delay between MS and radio access network (RAN) is t_{mr}. It is the time to send a message via the wireless link.
- The delay between MS and access router AR is t_s. This time includes the delay of the wireless link as well as the delay between RAN and AR.
- The delay between the old and new access point is t_{no}.
- The delay between the MS and AAA or CRRM entity is t_{mf}.
- The delay between the MS and HA is t_h. It is the time needed to deliver a message to the home network.
- The delay between the MS and correspondent node (CN) is t_{mc}.
- The delay between the home network of MS and CN is t_{hc}.

4.6.1 Signaling Delay of MIPv6-Based Handover

MIPv6-based handover includes the signaling shown in Figure 4.1, which includes the following signaling delays.

- The total delay of sending router solicitation signal is T_{rsol}

- The total delay of routing advertisement signal is T_{radv}
- The total delay of binding update signal is T_{bu}
- The total delay of binding acknowledgment signal is T_{ba}
- The total delay of handover test init signal is T_{hoti}
- The total delay of handover test signal is T_{hot}
- The total delay of care-of-address test init signal is T_{coati}
- The total delay of care-of-address test signal is T_{coat}

These delays are calculated as the sum of individual link delays involved in the signal along with the probability of retransmission and delay variation on the link. The simulation model generates a Gaussian noise model in generating effective link delay with a mean value shown in the model specified in Figure 4.7. All these signaling delays add up to the total handover delay involved in MIPv6-based handover procedure which is as follows.

$$T_{mipv6} = T_{rsol} + T_{radv} + (2 \times T_{bu}) + (2 \times T_{ba}) + T_{hoti} + T_{hot} + T_{coti} + T_{cot} \qquad \ldots\text{Eq. 4.1}$$

4.6.2 Signaling Delay of MCVHO

As most of the signaling delays are specified in MIPv6-based handover, the definition of the signaling delay for the same signal used in other handover schemes is not repeated. All other delays involved in MCVHO are specified as follows.

The total delay of vertical handover request and vertical handover reply in MCVHO is 2 x T_{mr}

- The total delay of authentication request signal is $T_{authreq}$
- The total delay of authentication reply signal is $T_{authrep}$
- The total delay of routing table update request signal is T_{rtureq}
- The total delay of routing table update reply signal is T_{rturep}

The total handover delay of an MCVHO scheme is given by Equation 4.2, which is the sum of all signaling delays involved in an MCVHO-based handover. The individual delay presented in this equation may have zero value because the signal is not sent as a part of the handover. For example, in the case of an intra-domain handover, the traditional MIPv6-based binding update signals are not sent. In addition, the value of authentication delay T_{mauth} is different when oPoA and nPoA are associated in different ways.

$T_{mcho} = 2 \text{ x } T_{mr} + T_{rsol} + T_{radv} + T_{mareq} + T_{mauth} + T_{marply} + T_{mrturq} + T_{mrturp}$ …Eq. 4.2

4.6.3 Signaling Delay of NCVHO

Similar to an MIPv6-based handover and MCVHO, the following signaling delays are involved in NCVHO.

- The total delay of vertical handover request signal is T_{mr}
- The total delay of profile request signal and profile reply signal is 2 x T_{rf}
- The total delay of authentication request signal is $T_{nauthreq}$
- The total delay of authentication procedure signal is T_{nareq}
- The total delay of authentication reply signal is $T_{naurply}$
- The total delay of routing table update request signal is T_{rturq}
- The total delay of routing table update reply signal is T_{rturp}
- The total delay of handover command signal is T_{mr}

The total delay of an NCVHO handover procedure is as follows.

$T_{ncho} = T_{mr} + 2 \text{ x } (T_{ra} + T_{af}) + T_{nareq} + T_{nauth} + T_{narply} + T_{nrturq} + T_{nrturp} + T_{mr}$ …Eq. 4.3

4.7 Link Delay Analysis

4.7.1 Effect of wireless link delay

For calculating the disruption time of a signal, t_{mr} is set to 10 ms as done in reference [30], considering the relatively low bandwidth of the wireless link. In addition, the other assumptions made in this model are $T_{mf} = t_{mr} + 2 = 12$ ms, $t_{no} = 5$ ms, and $t_s = 11$ ms. The delay attributed to the internet depends on various factors such as the protocol used, number of routers, and type of link used. The computation of delay in such a complex environment is very complex; therefore, one-way delay of internet is assumed to be constant with a value 100 ms [30]. Hence, $t_h = 112$ ms, $t_{mc} = 124$, and $t_{hc} = 114$ ms. In this section, the effect of delay on wireless link and delay on link between MS and HA on total handover delay of handover is studied; the results of the experiment are shown in the Figures 4.8 and 4.9.

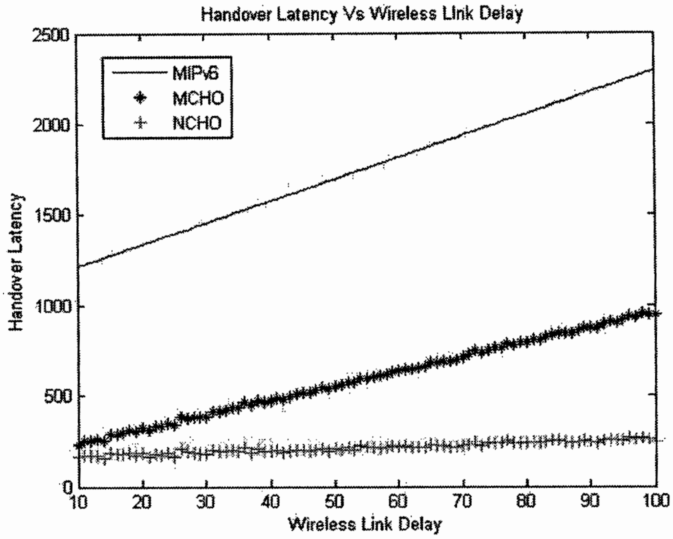


Figure 4.8 Handover Latency(ms) Vs Wireless Link Delay(ms)

The effect of wireless link delay on the overall handover latency is shown in Figure 4.8. All the figures in this thesis show average result of simulation executed for number of times to minimize randomness in the result. The figure shows this effect for the three schemes discussed in this chapter, namelyMIPv6-based handover, MCVHO, and NCVHO. It can be observed that, as the wireless link delay increases, the total handover delay of all handover schemes also increases. The handover delay of MCVHO and MIPv6-based handover are affected to a greater extent by increased delay on the wireless link as shown in Figure 4.8. According to Figure 4.1, all the signals of MIPv6 and according to Figures 4.4 and 4.5, all the signals of MCVHO traverse over the wireless link. As shown in Figure 4.6, only two signals of NCVHO use wireless link. In MCVHO and NCVHO, handover delay is also affected by the association between oPOA and nPOA. In this simulation, it is considered that 40% of the time the oPOA and nPOA have a direct association, 40% of the time they are connected through some AAA network, and 20% of the time the oPOA and nPOA have no association and authentication has to be done via a HA. This assumption does not affect MIPv6-based handover, because it always performs binding update via HA which results in a greater delay of handover. This can be observed in Figure 4.8 because the handover delay of MIPv6-based handover is highest throughout the simulation. Furthermore, the increase in an MIPv6-based handover delay is more than the increase in the handover delay of MCVHO and NCVHO.

4.7.2 Effect of Link delay between MS and HA

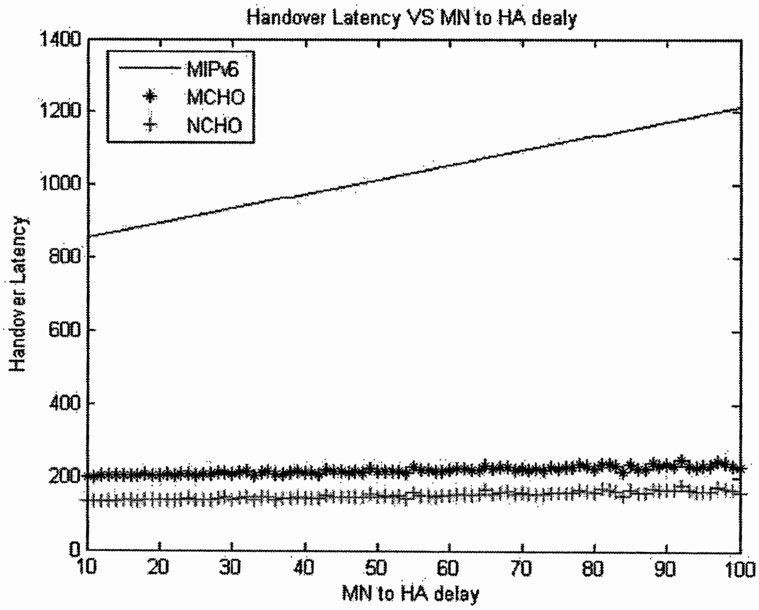


Figure 4.9 Handover Latency v MS to HA Link Delay

The effect of link delay between MS and HA on handover delay is shown in Figure 4.9. As described earlier, the association between oPOA and nPOA is assumed in the simulation. As a result, the effect of delay on the link between MS and HA has less effect on MCVHO and NCVHO. MIPv6-based handover is the most affected by this link delay because MIPv6-based handover always involves HA for authenticating the MS with nPOA.

4.7.3 Effect of FER

In this part of the simulation, the effect of size of messages exchanged between different nodes and effect of error and retransmission are considered. Delay is calculated as a function of Frame Error Rate (FER). The process of packet transmission is considered a random error process. The retransmission of a signal is based on the exponential back-off timer specified in reference [30]. The timer value T_r is described in the following equation.

$$\mathrm{Tr}(i) = 2^{i-1} \cdot \mathrm{Tr}(1). \quad \text{... (Eq.4.4)}$$

To calculate the delay in signal transmission, the total delay of the signal has to be calculated based on FER. The link delays between the various nodes of the network are calculated using the total delay of the signal on the wireless part. The simulation results of handover delay are shown in Figure 4.10.

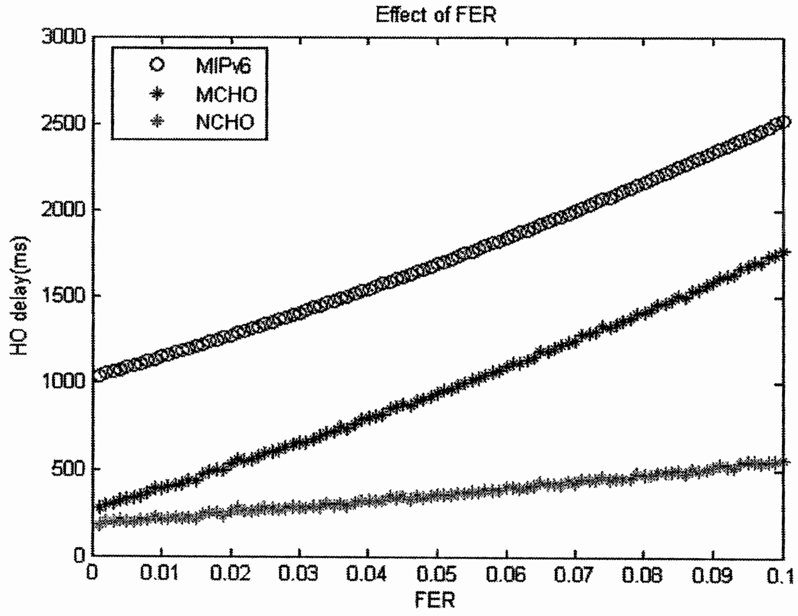


Figure 4.10 Effect of FER on Handover Latency

Figure 4.10 shows the effect of FER on handover delay corresponding to all the three handover schemes discussed in this chapter. The line on the top of the graph shows the handover delay of an MIPv6-based vertical handover when FER increases from 0.01 to 0.1. This result of MIPv6 is verified with the result presented in reference [30]. The delay associated with the MIPv6-based handover is the highest among all the three proposed schemes. From equation (4.1) we see that the signaling of an MIPv6-based handover has the longest delay because it involves the HA in the handover procedure. Handover using MIPv6 and MCVHO shows similar increases in handover delay with increasing FER. From equations 4.1 and 4.2, we see that both the schemes involve signaling on the wireless part of the network. Increase in FER leads to more errors and retransmission rate on the wireless link and results in increased link delay. This increased link delay in turn leads to increase in the overall handover delay. Equation 4.3 shows that the NCVHO scheme has very limited signaling done on the wireless part of the network. Therefore, the NCVHO scheme has little effect of increase in FER on the overall handover delay.

4.7.4 Effect of the Number of Users

The increase in the number of users is a more crucial parameter, which increases handover delay in NCVHO. This is because of the increase in the processing delay on the oPOA as the number of users increases. Each new user adds some fixed amount of procedural delay on oPOA which is caused by handover signaling. Increase in the number of users will also call for more scheduling which will also increase delay in the processing. However, in an MCVHO this delay does not increase that quickly because the handover decision is taken individually by each MS. The effect of an increase in the number of users is shown in Figure 4.11.

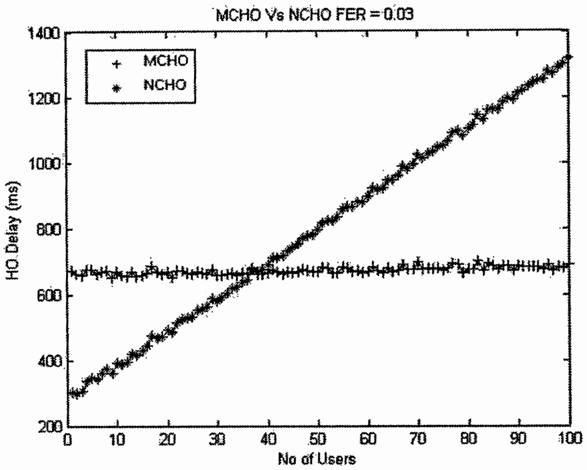


Figure 4.11 Effect of the Number of Users

The line marked with "*" signs shows the handover delay associated with NCVHO and the line with "+" signs shows handover delay associated with MCVHO. MCVHO has almost no effect of the increasing number of users whereas NCVHO is highly affected by the increase in the number of users. The simulation has been performed using an FER = 0.03 and the number of users increasing from 0 to 100. The graph shows that after the number of users exceeds 40 the delay of NCVHO becomes greater than that of MCVHO.

4.7.5 Effect of Number of Users and FER

To compare the effect of FER and the number of users together on MCVHO and NCVHO, we have simulated the handover procedure by varying FER and the number of users. Figure 4.12 shows the effect of FER and the number of users on the handover delay.

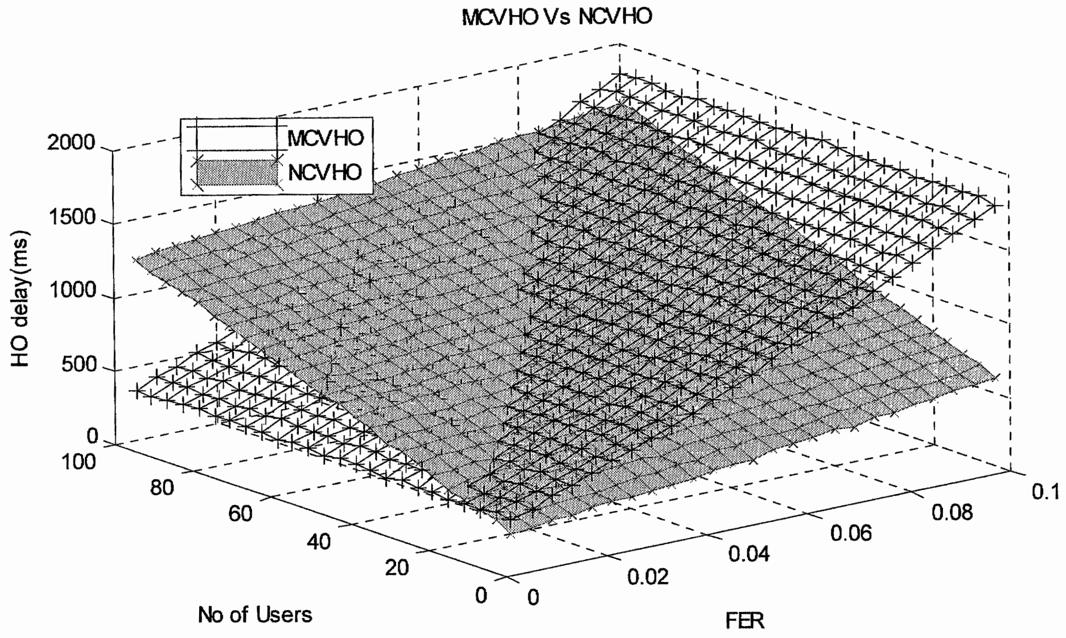


Figure 4.12 Effect of FER and the Number of Users on Handover Latency

It can be seen from Figure 4.12 that when the number of users is few, NCVHO performs better than MCVHO. As the number of users increases, the performance of NCVHO decreases compared to MCVHO. In contrast, as FER increases, the performance of MCVHO decreases. The point at which the performance of NCVHO gets worse than that of MCVHO moves higher with increasing FER. NCVHO has a centralized approach of handover decision making, which increases processing delay at node B. When a new user places a handover request, it adds up some fixed amount of processing which is needed to make handover decision. The increase in the number of users also increases processing delay because the time spent in scheduling user handover requests also increases. This computation increases handover delay as the response of node B decreases along with an increase in the number of users.

Authentication using direct association has already been proposed in reference [7]. The enhancement of authentication procedure using AAA architecture has also been proposed in reference [57]. These enhancements to perform fast authentication can be used in both MCVHO and NCVHO. These considerations help us to compare the performances of both the schemes and help us in deciding which scheme performs better in same situations.

4.8 Summary

In an MIPv6-based handover, authentication can be performed only via a HA, and this increases the delay of handover because the HA is connected with the oPoA and nPoA through some public network such as Internet. This delay in authentication can be reduced using direct association and AAA architecture as described earlier.

In the case of an MCVHO, signaling is done by the MS and it travels through wireless network. The signaling done on wireless network causes more retransmission and apparently more delay in the overall handover procedure because wireless network is slow and less reliable than wired network.

In NCVHO and MCVHO, fast authentication schemes using direct association and AAA architecture are used to enhance the delay in performance. In NCVHO, all the signaling is done by the oPoA instead of involving the MS; this increases the delay performance of the handover procedure. It should be noted that in MCVHO all the signaling is done by the MS which goes through the oPoA and nPoA. But in the case of NCVHO, instead of involving the MS, it is the oPoA which performs authentication, network selection, and RTU for the MS. This reduces signaling delay because the wireless part of the network is not involved in the signaling process.

The cost of reducing this link delay is paid with increased processing cost at oPOA. Previously, a network was evaluated individually by each MS, whereas now this procedure is carried out by oPOA and hence the processing has become centralized, thereby increasing delay in high-density cells. This issue needs to be resolved to achieve better delay performance of NCVHO in a high-density cell. This issue is covered in Chapter 6.

5 Network Selection Scheme

5.1 Vertical Handover and Policy Design:

To perform a vertical handover, the network employs certain policies on the basis of which it decides when, where, and how the vertical handover should be performed. Some of the parameters that are considered while performing a vertical handover are user preference, and context information. Horizontal handovers in homogeneous networks neither allow users to switch between different networks automatically nor is it concerned with the transfer of context information of users, which may include security, QoS requirements, and authentication parameters. A horizontal handover procedure is typically based on simple parameters such as signal strength and signal-to-noise ratio.

In future heterogeneous networks, the handover process from one network with particular type of radio network technology (RAT) to another network with another RAT will require policies and decision algorithms, which will take various factors such as cost, user preference, available services, and QoS into consideration. In the literature vertical handover decision schemes are proposed on the basis of different strategies such as neural networks, fuzzy logic [1], decision metric [29], analytical hierarchy model [28], and cost function [38]. These schemes also use various parameters and context information to make a vertical handover decision. In this chapter, these schemes and their performance based on the parameters used in making a vertical handover decision are analyzed extensively.

The aim of this chapter is to find an optimal set of parameters, a handover decision making scheme, and a handover control scheme to ensure that users will camp on the best target network. Making a right handover decision is critically important because an impropriate selection of a target network can result in frequent handovers in the network. This will cause more signaling, more congestion, and thus increased call dropping. The simulation results show how different sets of parameters used in a vertical handover decision making affect the amount of call dropping and the number of handovers. The simulation results also reflect the effect of a handover control scheme on the performance of vertical handover. In the following sections, different vertical handover decision making schemes are discussed.

5.2 Neural Network and Fuzzy Logic-based Scheme

A fuzzy logic-based handover scheme is discussed in [1]. In this scheme, a system makes a handover decision on the basis of parameters related to the sensitivity of the handover algorithm such as SIR, MS velocity, and RSS. A fuzzy logic-based system allows making a handover decision using different parameters in a complex system and includes fuzzifier, fuzzy interference system, fuzzy rule base, and fuzzy reasoning. The fuzzifier transforms crisp inputs into logistic values, which are inputs to the fuzzy logic system. These logistic values are used in fuzzy rule base, which allows making decisions based on the value of fuzzy variables. The fuzzy interference system maps this fuzzy rule base onto the fuzzy sets, and the output of fuzzy interference system is used by the defuzzifier, which generates crisp outputs from fuzzy results.

5.3 Decision Metric and Policy-Based Handover:

5.3.1 Decision Metric:

A decision metric-based vertical handover decision making scheme is proposed in [28]. This scheme works on a decision metric created using policies. The decision metric is a square metric, which contains decision parameters in the columns and their values in the corresponding rows. The decision parameters are service type, monetary cost, network condition, system status, and call parameters. Service type allows providing different aspects such as reliability, latency, and data rate. Monetary cost is calculated using user preference. Network condition shows whether the network has resources to host the call. System status provides information regarding network congestion. Call parameters include call requirements such as mobility requirements and bandwidth requirements.

Using all these parameters, the cost of candidate networks is calculated and a handover decision is made accordingly.

5.3.2 Policy Design:

While decision metric helps in deciding at which point to perform a handover, policy helps in deciding when to perform a handover. In Figure 5.1, a simple handover policy in a homogeneous network is shown. In this figure, *x*-axis shows the position and *y*-axis shows RSS. In this case, a handover can take place at any point from A to D. The exact point when a handover takes place depends on policy employed by the network. The policies are simple and efficient for the traditional handover. However, policies for a vertical handover have to be more complex and include more parameters to assist decision making in a complex environment.

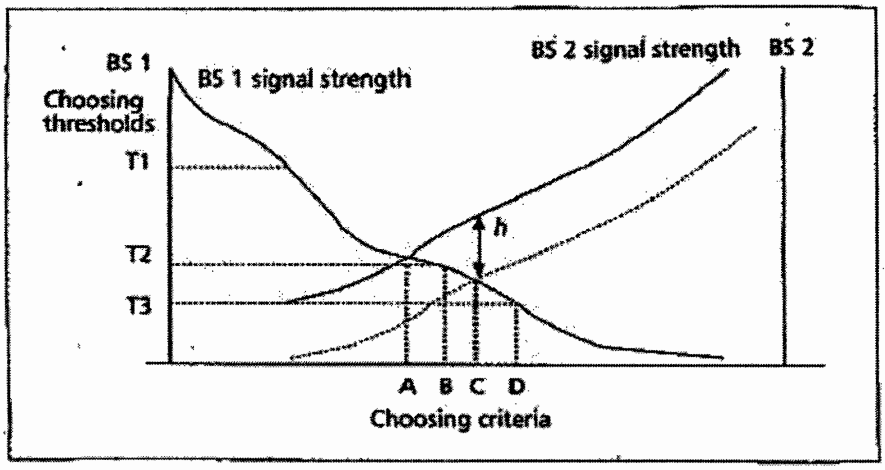


Figure 5.1: Traditional Handover Policies [28].

The dynamic vertical handover algorithm is one of the policies for making a handover decision in a complex environment. This algorithm uses the cost function for calculating the cost of performing a handover to the target network. The decision of when to perform a handover is made based on this cost; this decision-making scheme categorizes cost function parameters into three: QoS parameters, weighting parameters, and network elimination parameters. Calculation of this cost function is detailed later in this chapter.

5.4 Network Selection using Mathematical Modeling:

Several other techniques have been proposed by researchers for selecting a network while performing a handover; one among them is network selection based on mathematical modeling [30]. This scheme is a two-stage procedure, which determines the relative weights of the available networks on the basis of parameters related to the network and the user. The first stage of this procedure is the Analytical Hierarchy Process (AHP), which is defined as the procedure used for dividing a complex problem into a number of deciding factors and integrating the relative dominance of the factors with solution alternatives in order to find the optimal set [29]. The second stage of the procedure is the Gray Relational Analysis (GRA), which builds gray relationships between elements of two series in order to compare each member quantitatively [29]. AHP and GRA-based network selection schemes are shown in Figure 5.2 and is detailed in reference [29]. The WLAN environment has been given higher priority in this scheme because of the cost and bandwidth factors. Figure 5.2 shows the steps that include AHP and GRA.

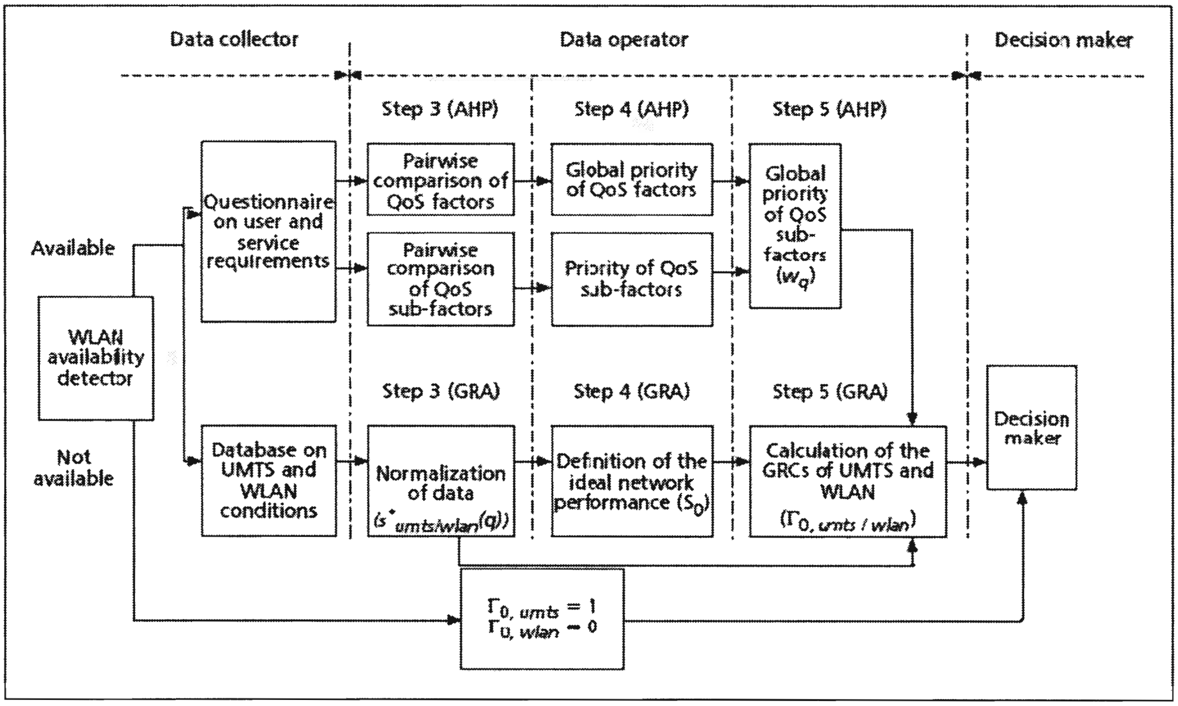


Figure 5.2: Handover Policy using AHP and GRA [29].

5.5 Cost Function-Based Handover Decision Scheme

The cost function-based handover decision scheme is proposed in References [28] and [38]. This scheme uses various parameters such as signal strength, available bandwidth, user required throughput, mobility requirements, and user's preference. In this scheme, a handover decision is made for interoperability between UMTS network and WLANs. The interoperability scheme is defined into two parts namely Initial User Assignment (IUA) and Inter-System Handover (ISH). The two schemes use different sets of parameters and allow users to select the network that is the most suitable for them under specified conditions. The schemes employ cost function created based on user's preference function to evaluate the parameters which are input to the cost function. The evaluation to make a handover decision is based on policies. Both preference function and cost function are defined and discussed in the following paragraphs. In addition, simulation results for evaluation of such schemes are given. The simulation developed in this thesis to make a handover decision uses the cost function-based handover selection scheme to make a handover decision.

The cost function is composed of various weighted parameters such as available bandwidth supported by the network and maximum transmit power at user terminal. The policy based on these parameters is defined on the basis of preference function. This preference function is defined as the preference for a mobile terminal to camp on UMTS or WLAN in the case of interoperability between UMTS and WLAN. Let X be the parameter and $P_f(X)$ be the preference function for parameter X; then, the preference function $P_f(X)$ can be defined as shown in the graph (Figure 5.4). As the value of X changes, the preferred network changes. The value of P_f for any parameter always remains between 0 and 1. As this value approaches 0, WLAN is preferred and as the value reaches 1 UMTS is preferred.

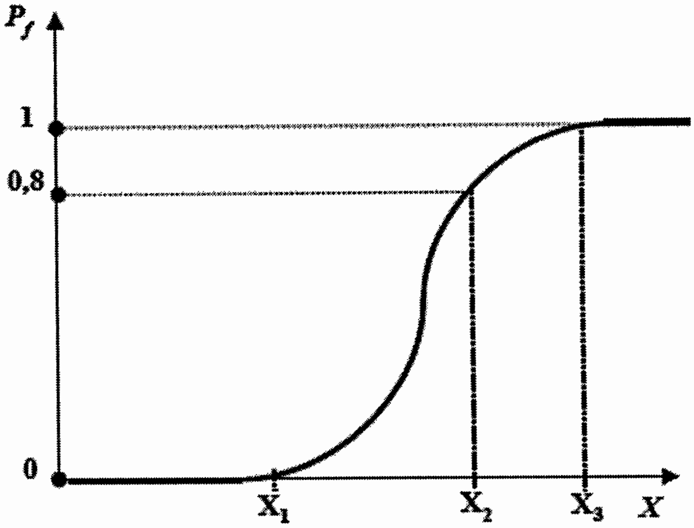


Figure 5.4 Preference Function Computations [38]

The criterion X is not required to have a continuous value and may have two or three discrete values. For example, X may take one of the three values X_1, X_2, or X_3. At each point, the value of preference function is 0, 0.8, and 1, respectively. Computation of the preference function is based on the following equation [38].

$$P_f(X) = \alpha_{0,X}\overline{P_0(X)} + \alpha_{1,X}P_1(X) \text{ with } \alpha_{0,X} + \alpha_{1,X} = 1 \qquad (5.1)$$

Where, $P_0(X)$ is the probability of using criterion X for choosing WLAN, $\overline{P_0(X)}$ is complement of $P_0(X) = (1 - P_0(X))$ and $P_1(X)$ is the probability of criterion X for choosing UMTS. $\alpha_{0,x}$ is the weight corresponding to the probability $\overline{P_0(X)}$; $\alpha_{1,x}$ is the weight corresponding to the probability $P_1(X)$. Computation of the preference function is detailed in reference [38]. Use of this function allows the prioritizing of different parameters to make a handover based on network and user preferences. In this computation, the weight of the parameters can also be different for various network and user conditions.

To compute cost function, different parameters are weighted and then multiplied with the preference function value. The sum of the weighted parameters is between 0 and 1. If the evaluation of cost function results in choosing a different network other than the current one, a vertical handover will take place, whereas if the evaluation results in the same current network an intra-system handover will occur. The cost function criteria for IUA and ISH are based on different parameters. In general, cost function can be represented by the following equation.

$$\mathrm{CF} = \sum_{X} \beta_X P_f(X) \text{ with } \sum_{X} \beta_X = 1 \qquad \text{... Eq. 5.2}$$

where, β_x is the weight of criterion X. The preference function computed for each criterion are weighted and then summed up to compute the cost function value. The total of all weights is equal to 1. Weighting of the factors is bound by the following rules [38].

- $\exists X / P_f(X) = 0 \wedge \forall Y \neq X, P_f(Y) \neq 1 \Rightarrow \beta_X = 1 \wedge \beta_Y = 0$
- $\exists X / P_f(X) = 1 \wedge \forall Y \neq X, P_f(Y) \neq 0 \Rightarrow \beta_X = 1 \wedge \beta_Y = 0$
- $\exists X / P_f(X_1) = 0 \wedge \exists X_2 / P_f(X_2) = 1 \Rightarrow \forall X, \beta_X = 1/N$
- $\forall X, P_f(X) \in \lfloor 0,1 \rfloor \Rightarrow \forall X, \beta_X = \lambda$

where, N is number of monitored criteria and λ is a scalar.

Given that the value of preference function is between 0 and 1, and the sum of weight β_X is also 1, the value of cost function is also between 0 and 1. To achieve higher performance of the handover decision process, weight tuning is performed. This involves changing the weights of the parameters dynamically. One of the weighting changing techniques proposed in the literature is a closed loop system, which has performance analysis as a feedback system for weighting the parameter [38]. In this system performance analysis of the current state is given as input to the next state so that the weight of parameters can be changed to enhance the performance.

A cost function-based handover selection algorithm is described in [36]. The advantage of a cost function-based network selection process is that it tunes parameters dynamically and reflects the current requirements of the user. The selection of the parameter set can also be done adaptively; however, this may increase computational load. This method has described sets of parameters for IUA and ISH, which are discussed in the following sections of this chapter. This experiment is aimed to find the optimal set of parameters and analyze performance of the cost function-based handover decision making as well as to make sure that handover process achieves the best possible service conditions for the user. This method allows giving dynamic preference to the network based on current requirements. The performance analysis is done in terms of the number of handovers, call dropping probability, user throughput, and network load balancing.

5.6 Network Selection using Cost Function

This section presents a set of simulation results for cost function based network selection algorithms. The simulator was developed using MATLAB running on the UTS Faculty of Engineering's Cluster. The first simulation was designed for analyzing the performance of the cost function. This simulation takes four parameters for the IUA process and four criteria for making the ISH decision. The parameters used for performing IUA and ISH are different. The parameters used to perform IUA are terminal type, traffic characteristic, current user speed, and user preference, whereas those used to perform ISH are mobility requirement (based on type of equipment user is using), link quality, target cell load, and user required throughput.

The aim of performing an IUA is to assign a user to a particular network that is the most appropriate to its terminal type and type of services the user will likely access. The parameters that are used to perform IUA help in predicting the service pattern of the user. In a similar way, the parameters used for performing an ISH simulation help to decide whether there is a need for performing a vertical handover; as mentioned earlier, the ISH parameters are mobility requirement, link quality, target cell load, and user throughput.

Mobility requirement highly depends on the mobility pattern of the user and the type of terminal that is used by the user to access services. If the user has a highly-mobile device such as mobile handset or PDA, then it is likely that the mobility requirements of the user are high whereas if the user has a laptop the mobility requirements are usually low. The link quality of a target PoA can be decided by the received signal strength of the beacons or pilot signals received from that PoA. Cell load information of the target cell is available from the CRRM entity, and finally user throughput requirement is available by the type of service accessed by the user.

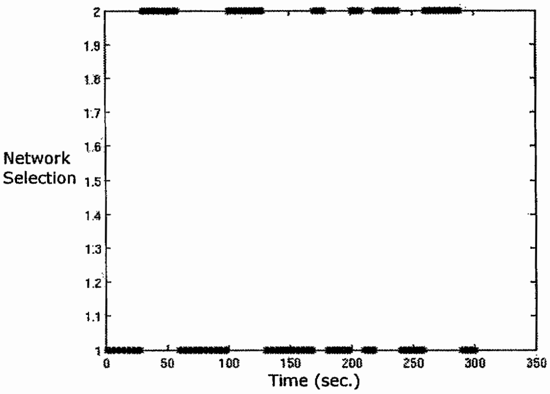


Figure 5.5: Low mobility mix calls

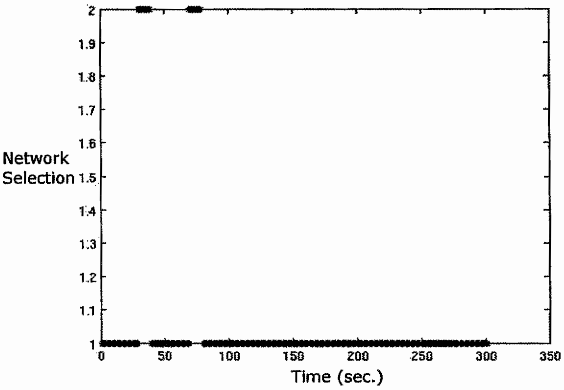


Figure 5.6: Low mobility data calls

The simulation is done to check network selection pattern using cost function under different traffic conditions. In Figure 5.5, network selection is performed when the user has a low mobility terminal such as laptop and the user is making both types of calls, namely, voice calls, which are more appropriate to UMTS, and data calls, which are more appropriate to WLAN. The dots at value 1 represent that WLAN is selected and those at value 2 represent that the selected network is UMTS. The *x*-axis represents time in seconds.

The terminal in Figure 5.6 is for making data calls only. The results show that mixed call traffic has a mix selection pattern, where nearly half of the time, WLAN is preferred and the rest of the time, the terminal is connected with UMTS network. In the case of data calls, the UMTS is selected only two times. Therefore, it can be inferred that network selection reflects the pattern of user throughput as well as mobility requirements. WLAN is a better choice for low mobility and high data rate, and so low mobility and data calls are assigned to WLAN, whereas UMTS is the preferred network for voice calls because it provides reliable service with a wide coverage area.

In Figures 5.7 and 5.8, simulation is done for a high-mobility user. In Figure 5.7, the user is making mixed calls, whereas in Figure 5.8 the user is making voice calls only. The selection pattern on the basis of cost function reflects user mobility and traffic requirements similar to Figures 5.5 and 5.6. In Figure 5.7, the user is connected with WLAN only few times. This is due to the fact that as the speed of the user goes beyond 5 kmph, WLAN is not preferred because the user is likely to move out of coverage very soon. Also, in Figure 5.8, voice calls are assigned to the UMTS network.

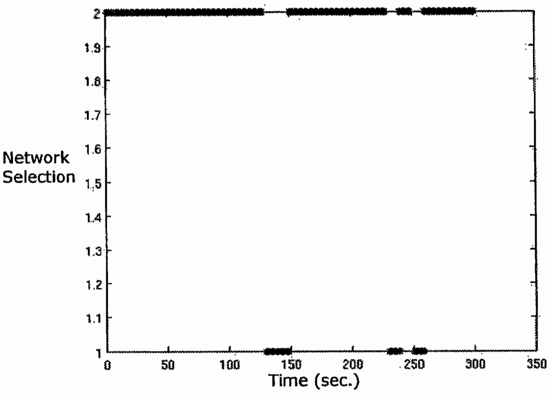


Figure 5.7: High mobility mix calls

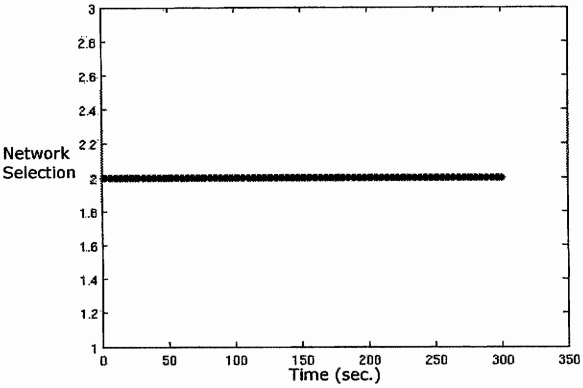


Figure 5.8: High mobility voice calls

In the earlier section, the use of cost function for selecting the network appropriate to user preference and network conditions was discussed. In the case of an MCVHO, the information has to be available at the mobile node for the decision to be made, whereas in the case of an NCVHO the information has to be available at the network node.

The study of network presented in reference [25] shows that some information can be calculated by the network only, such as network load and user speed. In order to use this information for enhancing handover decision making in MCVHO, this information has to be passed from the network to the mobile node. This transfer of information can again increase signaling on the wireless part of the network as well as introduce further delay in handover. Therefore, to make an effective handover decision in MCVHO, the transfer of information needs to be minimized as much as possible. Similarly, in the case of NCVHO, the transfer of information from the mobile node to the network should be kept at minimal.

5.7 Simulation

The simulation is designed in such a way that both the schemes can be tested in the common model. A network selection procedure can have many types of networks involved in it. The simulation framework used in this simulation includes two overlaid networks, namely UMTS and WLAN. These two networks differ widely in their characteristics. UMTS provides limited bandwidth such as up to 384 kbps over a large geographical area, whereas WLAN provides very high bandwidth such as 100 Mbps but for a limited coverage area. UMTS allows a user to access services seamlessly throughout its wide coverage area with a high level of reliability but has limited bandwidth and high cost. WLAN, on the other hand, has a limited coverage area often covering a building or a small highly-populated area known as hotspot. This type of connection is ideal for data services such as web browsing and video streaming. A network topology used for the simulation is shown in Figure 5.9. There are four UMTS cells wrapped around along the line drawn. At the border of these cells are four hotspots as shown in the figure. The users move along the area covered by a square shown in Figure 5.9.

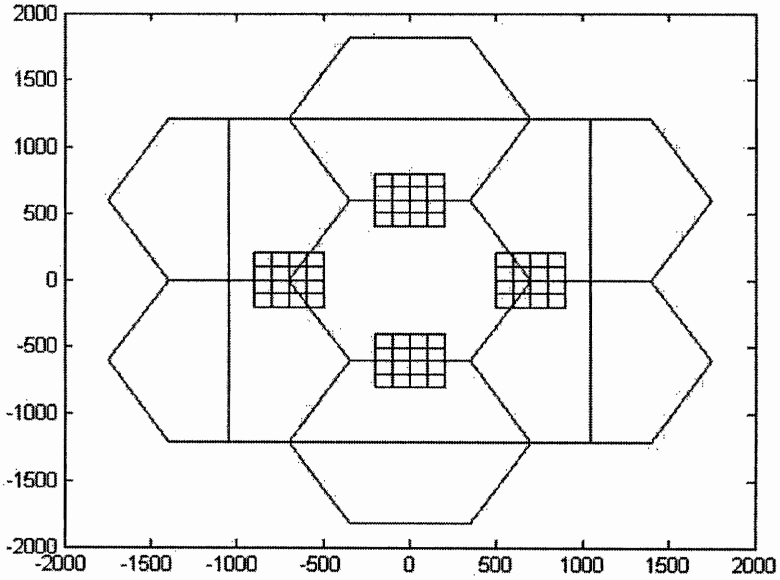


Figure 5.9 Network topology in the simulation

The complete simulation procedure is shown as a flowchart in Figure 5.10. Users are generated in a particular area. A random-walk mobility model [38] are employed to generate users' movement across the network. Two types of network are present in the simulation scenario, namely UMTS and WLAN. The path loss models are considered for radio propagation. Hotspot areas are indoor areas for all the inner cells. The outer cell of the hotspot covers both indoor and outdoor coverage areas. The path loss models of radio transmission for indoor and outdoor are given by the Equation 5.3 and Equation 5.4.

In the case of indoor users, the path loss model is based on a one-slop model where the slop factor $n = 2.4$, and a zero-mean lognormal fading with standard deviation of 7dB is added to the path loss model to include the shadowing effect. The equation of the path loss becomes as follows [38]. In the following equation, d is distance in km and S_1 is a zero mean Gaussian distributed random variable with 7 dB standard deviation.

$$P_L = 46.7 + 24 \log_{10}(d) + S_1 \quad \ldots \text{Eq. 5.3}$$

The linear attenuation model of path loss is used for outdoor users [38]. The path loss includes free-space path loss and additional loss caused by penetration obstacles in the environment. The lognormal fading with standard deviation of 10dB is added to include the effect of shadowing. The path loss equation is shown as follows.

$$P_L = 46.7 + 20 \log_{10}(d) + a.d + S_2 \quad \ldots \text{Eq. 5.4}$$

In the case of a UMTS system, the following path loss model is used, where R is the distance in km.

$$L = 128.1 + 37.6 \log_{10}(R), \text{ R in Km} \quad \ldots \text{Eq.5.5}$$

Figure 5.10 shows the complete procedure of simulation. In the first part, users are generated at the start of the simulation. IUA is performed at the start of simulation and users are allocated to the appropriate network according to the cost function calculated for IUA.

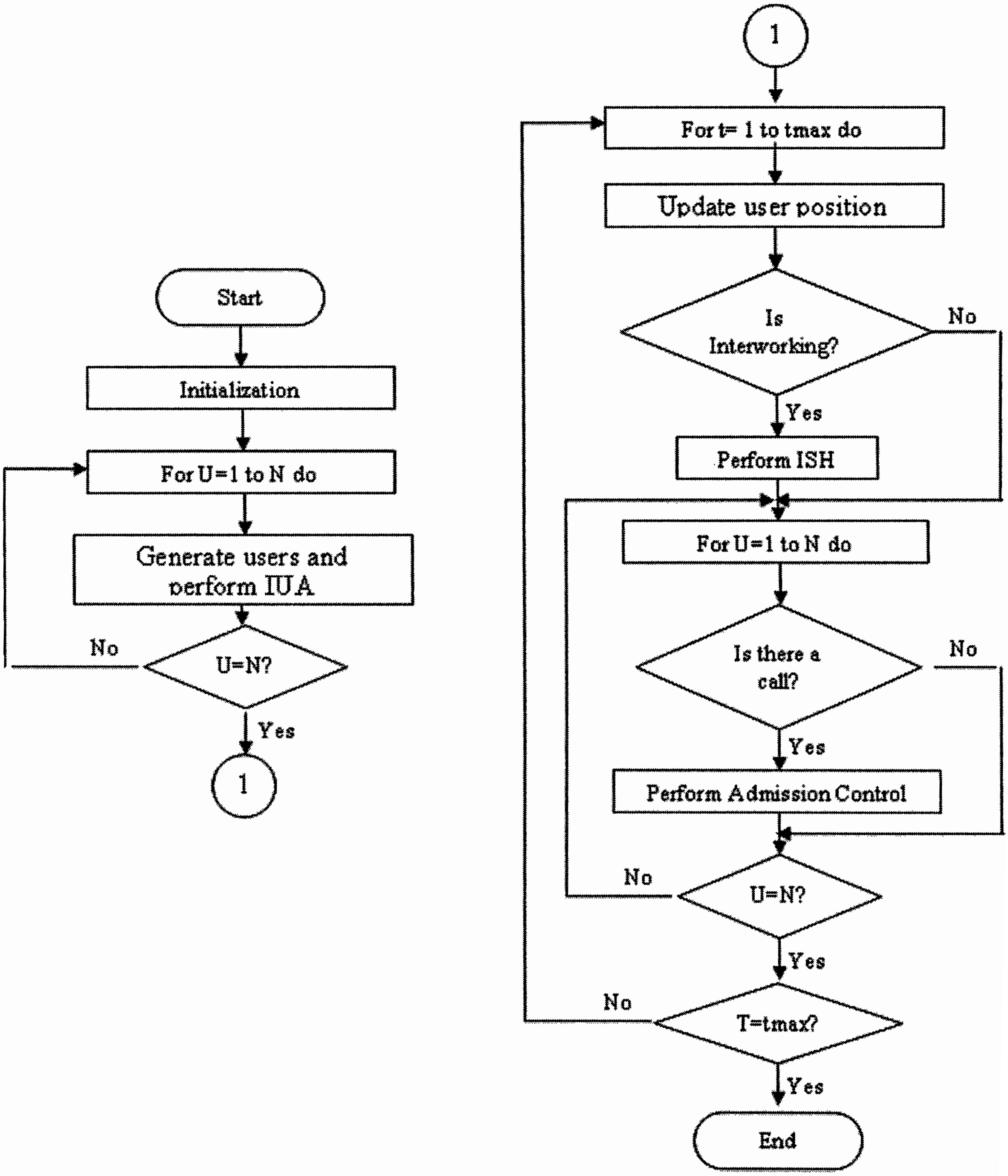


Figure 5.10 Flowchart of simulation

For each 1 second time interval, users are moved across the network using the random-walk mobility model. User traffic is generated using a traffic algorithm based on predefined call arrival probability and mean call holding time. Call arrival is a Poisson distribution based on lognormal random number generation. Admission control is based on the simple load threshold of UMTS. The ISH algorithm is based on the cost function using the four parameters specified previously. The simulation period is 10,000 sec. During this period, parameters such as call traffic, handover traffic, call blocking, and call dropping are registered for performance analysis. As specified earlier, simulation runs for both NCVHO and MCVHO schemes. To observe the effect of handover control scheme correctly, the same mobility pattern of users and traffic generation is used for both the schemes.

Performance analysis done on both the schemes includes tracing of user distribution among both the networks, tracing the traffic pattern of the simulation, handover calls, call blocking, and call dropping.

This simulation generates performance parameters such as number of calls blocked, number of handovers, and call dropping. The simulation also generates results to check how many call attempts have been successfully made in the favored network. An extensive number of handover calls implies more signaling on the radio part. As the number of call blocking and call dropping increases, it leads to user dissatisfaction and business loss.

As shown in the algorithm, in the first step, network topology and users are generated. The network performs IUA in order to allocate users to the appropriate network. The users will stay in that network until they are handed over to another network using ISH algorithm. The selection procedure is based on cost function as described previously. The results of the simulation are discussed in the following section.

5.8 Results

In Figures 5.11–5.16 the tracing of user distribution for the entire simulation period is shown. The graphs show the number of users in both the networks at each time interval. In Figure 5.11, the tracing of user distribution in the case of MCVHO is shown.

The broken line at the top of the graph represents the total number of active users in both the networks. The line represented by the "*"shows the number of users that are currently attached with a WLAN network. The line represented by "+"shows the number of active users in WLAN. In a similar way, the line represented by "square" shows the number of users attached with UMTS network and the bottom most line with "x" shows the number of active users in UMTS. The graph shows IUA in both the networks, at time = 0, which is nearly 50 for both the networks. As time goes on, more users approach WLAN using the selection criteria used in MCVHO and comparatively less number of users are assigned to UMTS. As the time of simulation goes higher, both the networks stabilize, and user distribution becomes nearly equal in both the networks. Increase and decrease in the number of active users in a network can result from call end, call drop, or handover to other network.

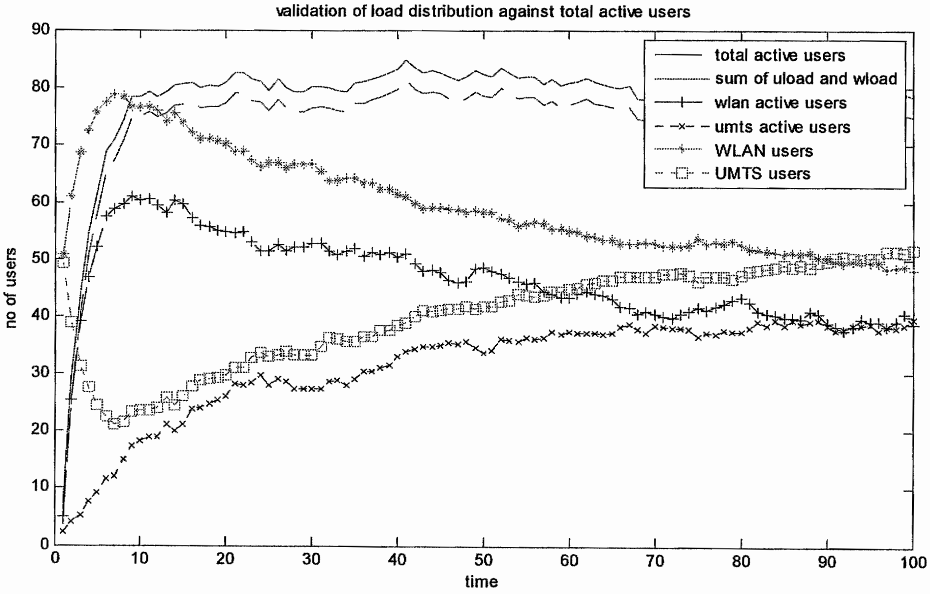


Figure 5.11 User Tracing for MCVHO

Figure 5.12 traces user distribution for the NCVHO algorithm similar to Figure 5.11. The graph shows the IUA for both the networks at time = 0. The IUA is performed using the same parameters for both networks. The user distribution and traffic models are the same for both the algorithms to make sure that performance is dependent only on the handover control scheme. The initial user distribution is also similar for both the schemes. In Figure 5.12, we find that user distribution among the networks is more balanced than in

Figure 5.11. This shows that the NCVHO scheme is more effective in balancing user distribution among different networks.

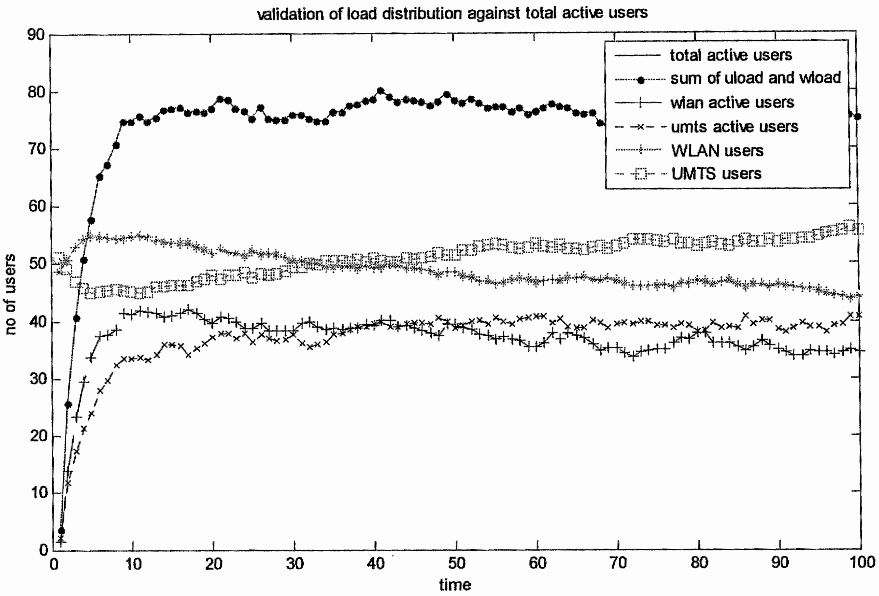


Figure 5.12 User Tracing for NCVHO

Since the network load information is available only with CRRM and passing this information to a mobile station is not feasible, the parameters employed by the MCVHO algorithm in selecting a network do not include load information. It is clear that the NCVHO allows the consideration of network congestion a handover decision making. It was noticed that the load-balancing property of NCVHO initially caused a ping-pong effect and resulted in more number of handovers. To avoid this effect, hysteresis for load balancing was introduced in the NCVHO scheme. The modification in the preference function considered a margin of 5% in the load distribution to take network load balancing into account for making a handover decision.

As described previously in this chapter, although handover decision algorithms are discussed in many literatures, the source of information utilization is not specified clearly. In the simulation performed here, the source of information and its utilization are considered and tested to find optimal weights of performance parameters.

5.9 Effect of Handover Parameters on Performance

To check the effect of handover decision-making parameters, different weights of parameters are used in different experiments. The results of these experiments are shown in Figures 5.13–5.18. The effect of changing the weight of parameters is also discussed in this section.

Figures 5.13 and 5.14 show the trace of user distribution among networks throughout the simulation using fixed, equal weights for all the four parameters. Figure 5.13 represents trace of user distribution among networks when MCVHO is used as a handover scheme, whereas Figure 5.14 is also a similar trace when NCVHO is used as a handover scheme.

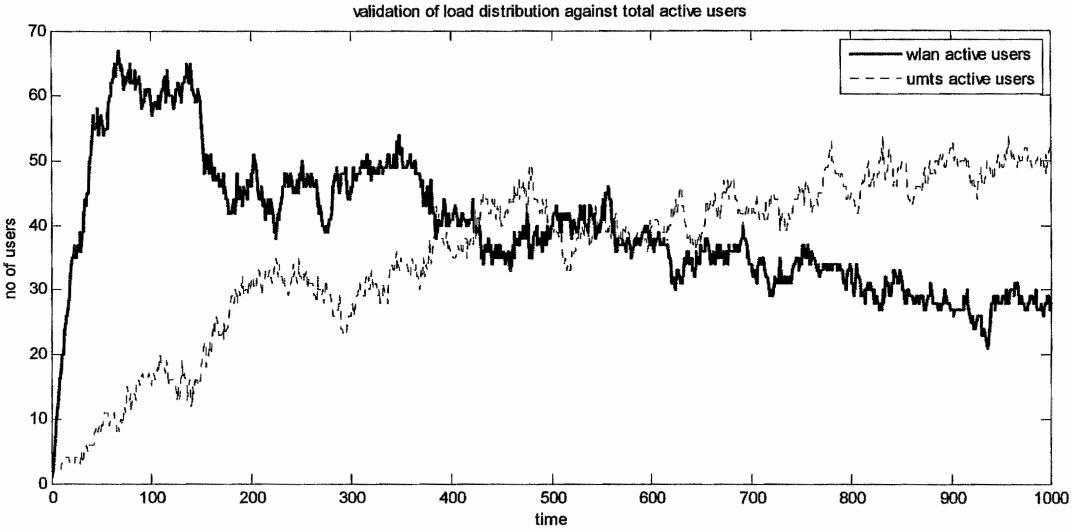


Figure 5.13 User trace of UMTS and WLAN for MCVHO using equal weight of parameters

The trace in both figures shows that user distribution in UMTS and WLAN keeps changing over the entire period. In the initial stage, more users are connected with WLAN. However, user distribution keeps moving toward the UMTS network along with the simulation period. As depicted in Figure 5.13 and 5.14, handover schemes have no control on user distribution between different networks. Figure 5.14 indicates that the NCVHO has slightly more control over user distribution among networks although the difference is insignificant.

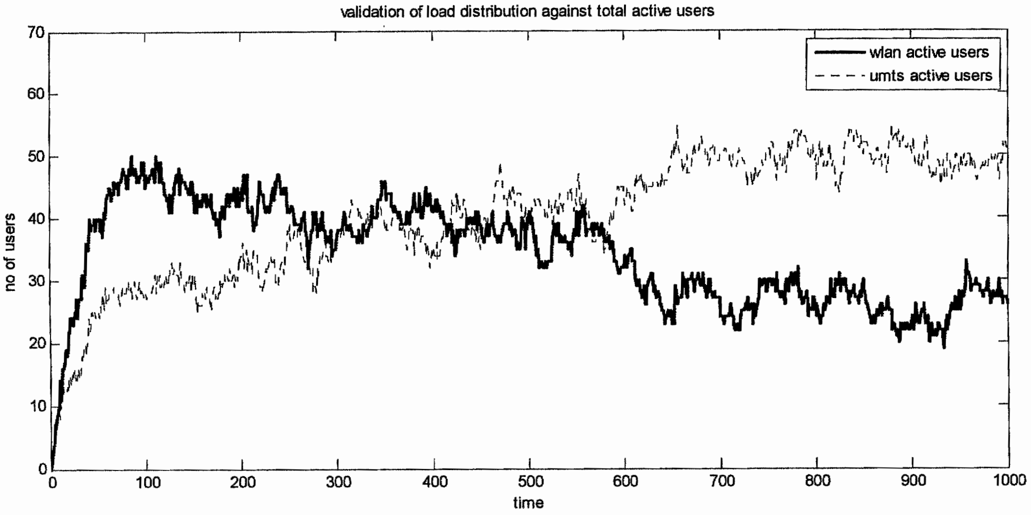


Figure 5.14 User trace of UMTS and WLAN for NCVHO using equal weight of parameters

In Figures 5.15 and 5.16, the trace of user distribution in the networks is shown when the weight of the parameters is tuned and different parameters are given different weights, which remain fixed throughout the simulation.

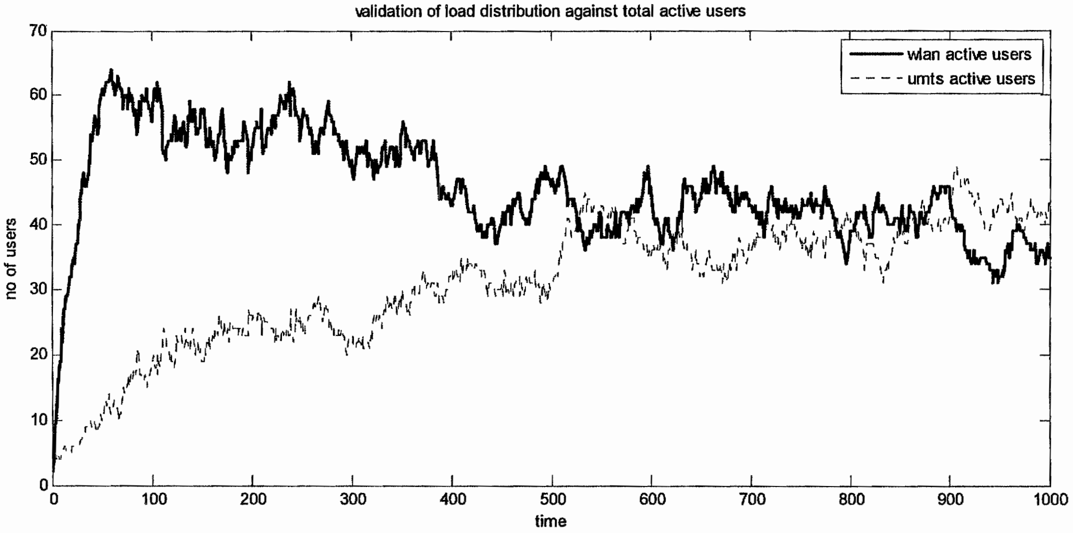


Figure 5.15 User trace of UMTS and WLAN for MCVHO using different weight of parameters

Figure 5.15 shows user distribution when MCVHO is used and Figure 5.16, when NCVHO is used as the handover scheme. Figure 5.15 indicates that there is a huge difference in the user distribution initially in MCVHO. However, along with time, user distribution becomes similar in both the networks.

In contrast, in the case of NCVHO in Figure 5.16, there is less difference in user distribution initially. However, this difference remains similar throughout the simulation. It is concluded from Figure 5.15 and 5.16 that different weights of the parameters have improved the performance of both the schemes. In addition, the performance of NCVHO is better than that of MCVHO because the NCVHO has more control over user distribution among the networks.

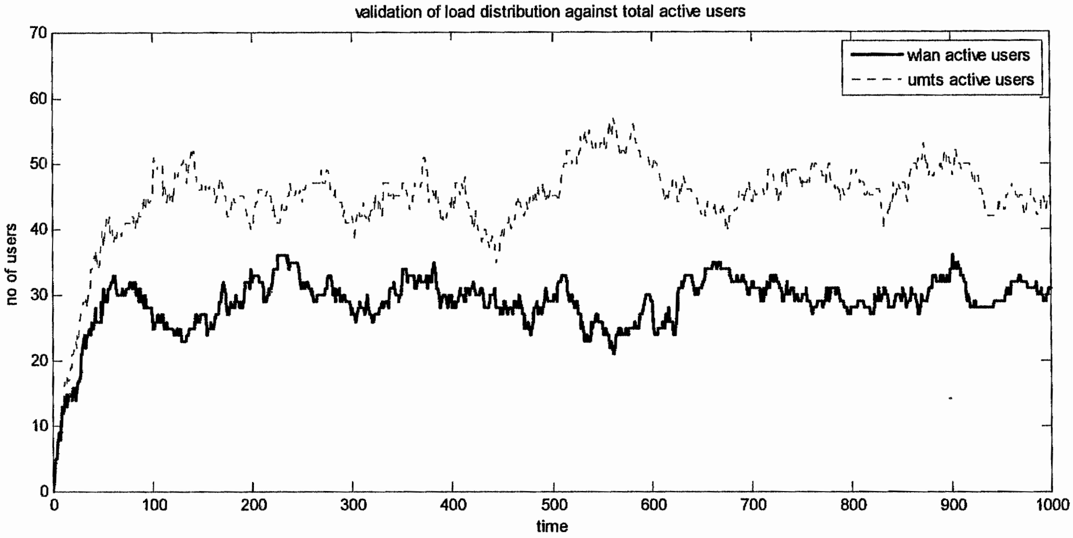


Figure 5.16 User trace of UMTS and WLAN for NCVHO using different weight of parameters

In Figures 5.17 and 5.18, weight tuning is applied on the go. The preference functions of the parameters are designed in such a way that handover decision making represents the current need of the user session. The results show that MCVHO has a stabilization pattern similar to the previous set of parameters under this situation, but the stabilization point has come earlier. The scheme has brought point of equal user distribution at time = 400, which is better than the previous one because this indicates that call dropping will decrease as a result of equal distribution. In the case of NCVHO also, difference exists in the user distribution still but there is less variation in the number of users. In Figure 5.18, more users are assigned to the UMTS network throughout the simulation period. This is because more than half of the users have speed >5kmph which are not assigned to WLAN by the NCVHO scheme. Furthermore, from the users with speed <5kmph, voice users will also have more preference to the UMTS network. These reasons justify more user penetration in the UMTS network in the case of the NCVHO scheme.

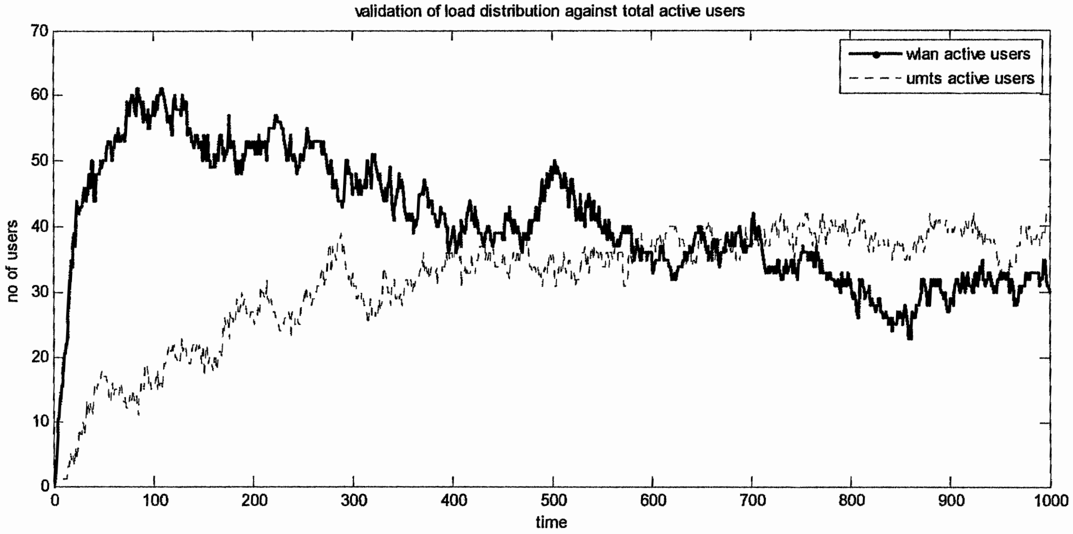


Figure 5.17 User trace of UMTS and WLAN for MCVHO using variable weight of parameters

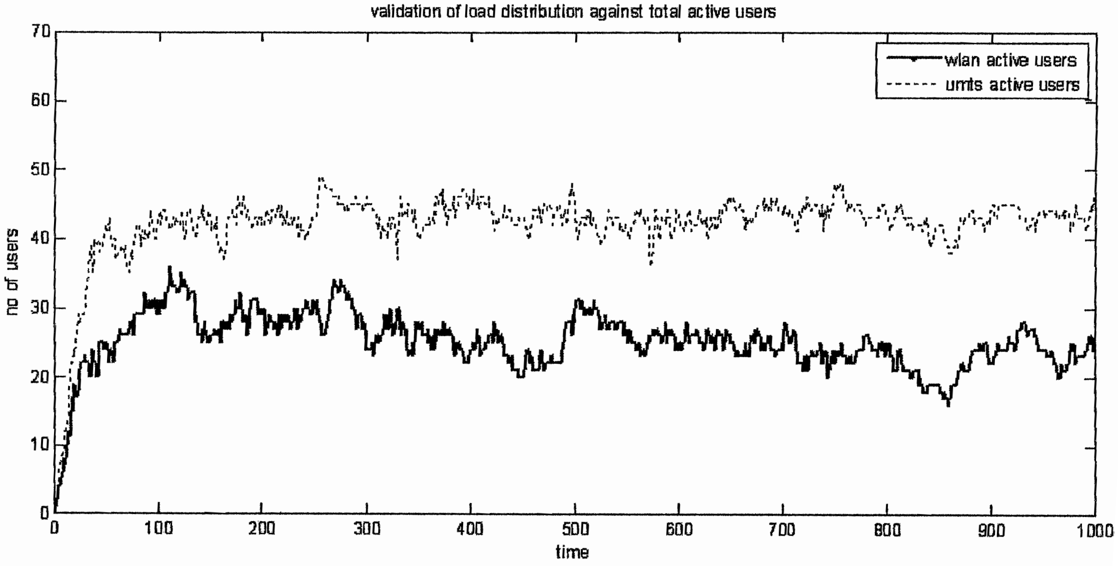


Figure 5.18 User trace of UMTS and WLAN for NCVHO using variable weight of parameters

5.10 Performance in terms of Number of Handovers

In the next step of simulation, after optimal weight tuning for both the handover schemes, MCVHO and NCVHO are compared in terms of the number of handovers, call dropping, and user satisfaction in order to observe which algorithm performs better.

Figure 5.19 shows the mean number of handovers that occurred during the entire simulation period with the number of users varied between 50 and 150. To check whether the results are consistent, the simulation is run 100 times and then the mean is taken for each user number. For each user number, both the handover schemes are tested over the same user distribution pattern, user mobility pattern, and traffic pattern. This procedure is repeated for each user number and the result is plotted as a graph (Figure 5.19). The information collected during the simulation includes the number of handovers occurred, number of calls generated, number of calls blocked, number of calls dropped, and number of successful attempts in a favored network.

The simulation is validated by monitoring user position, traffic generation, and handover decision making for a small period by going step-by-step. These checks are performed at various time points and under various conditions to monitor variations in the behavior of the handover schemes.

It can be seen from Figure 5.19 that less number of handovers has occurred in the case of NCVHO. In order to lead network to the saturation point, amount of call arrival is kept higher than the usual call arrival rate. This indicates that the network residency time of users is better when NCVHO is used as the handover scheme. Network residency time is the period for which a user is connected with the same network. This period increases when the user is assigned to a more appropriate network, which decreases the probability of handover to another network. This decrease in the number of vertical handovers reduces signaling on the network and consumption of network resources. As a result, the user obtains more satisfactory services.

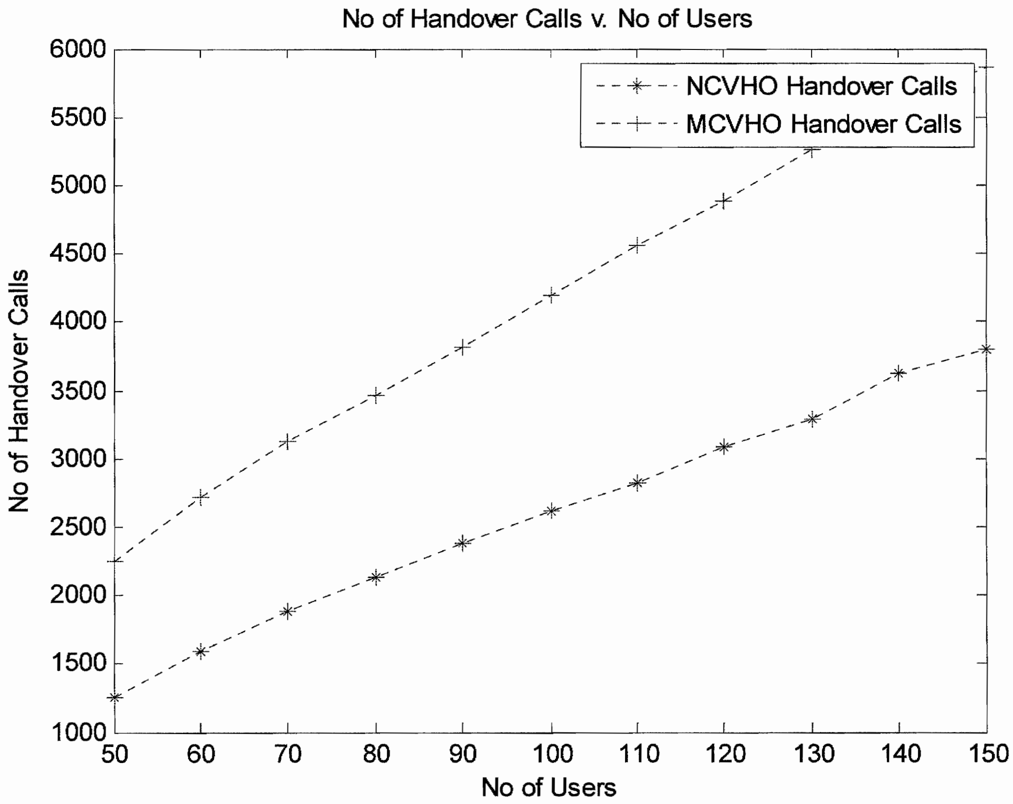


Figure 5.19 Number of Handover in the Simulation

5.11 Performance in terms of Call Dropping

In addition to the number of handovers, call dropping is another important parameter in the performance analysis of a handover scheme. Call dropping occurs when a call handover fails. In our simulation, this failure can occur because of two reasons: the handover attempt fails and the current network is not capable anymore to continue the call or the user moves out of coverage area of one network and is not accepted by the new network. In Figure 5.20, the call dropping that occurred during the entire simulation period is shown.

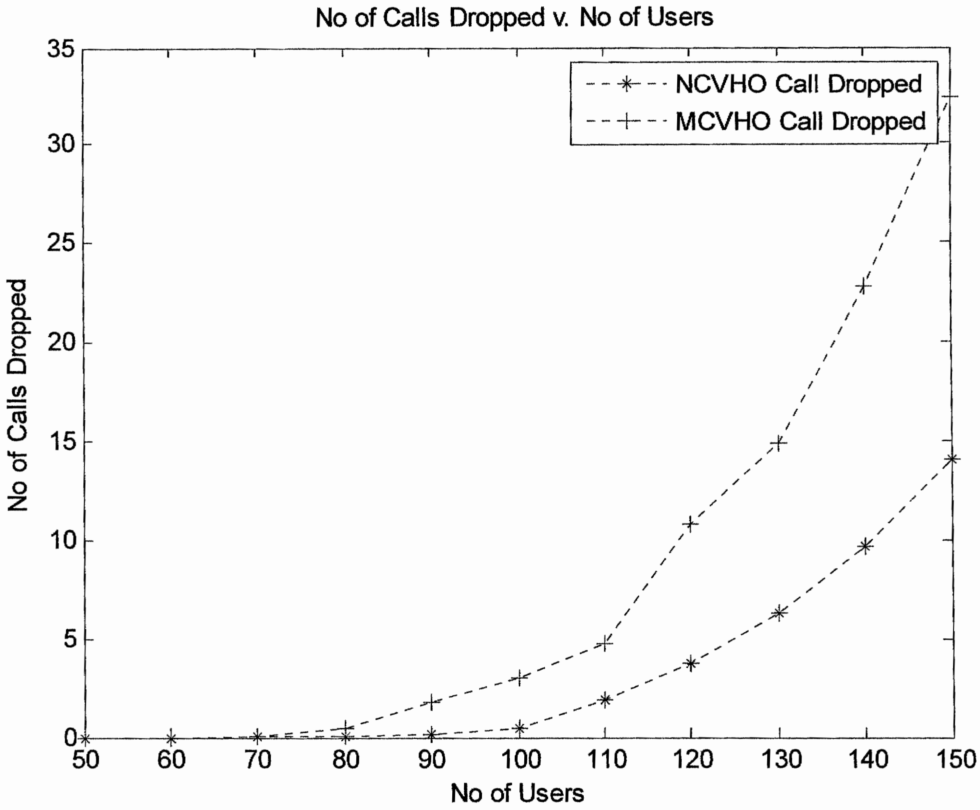


Figure 5.20 Number of Calls Dropped Vs Number of Users

In this case also, the mean number of call dropping in the simulation is shown. The simulation result is average based on 100 independent sets of simulation to obtain reliable results. Call dropping is a serious problem in the network. Users feel very dissatisfied when calls are dropped and therefore most of the networks take good care to ensure that call dropping is minimized. The mean number of call dropping for MCVHO and NCVHO is shown in Figure 5.20. It can be observed that NCVHO has a less degree of call dropping for all user densities compared to MCVHO. Call dropping for both the schemes have exponential growth when the number of users increase. However, the growth in the case of NCVHO is less than in the case of MCVHO. NCVHO takes network congestion and user mobility pattern in the consideration. Therefore, NCVHO achieves less call dropping than MCVHO.

5.12 Performance in terms of User Satisfaction

Another important performance criterion involved in a vertical handover is user satisfaction. This term refers to the performance criterion that the user is always connected to the best available network. The best available network is the network that can provide the best service with reasonable data rates and QoS. In this work, user satisfaction is defined as follows:

A user is considered satisfied if the user is connected to the preferred network for the type of service he/she is currently accessing.

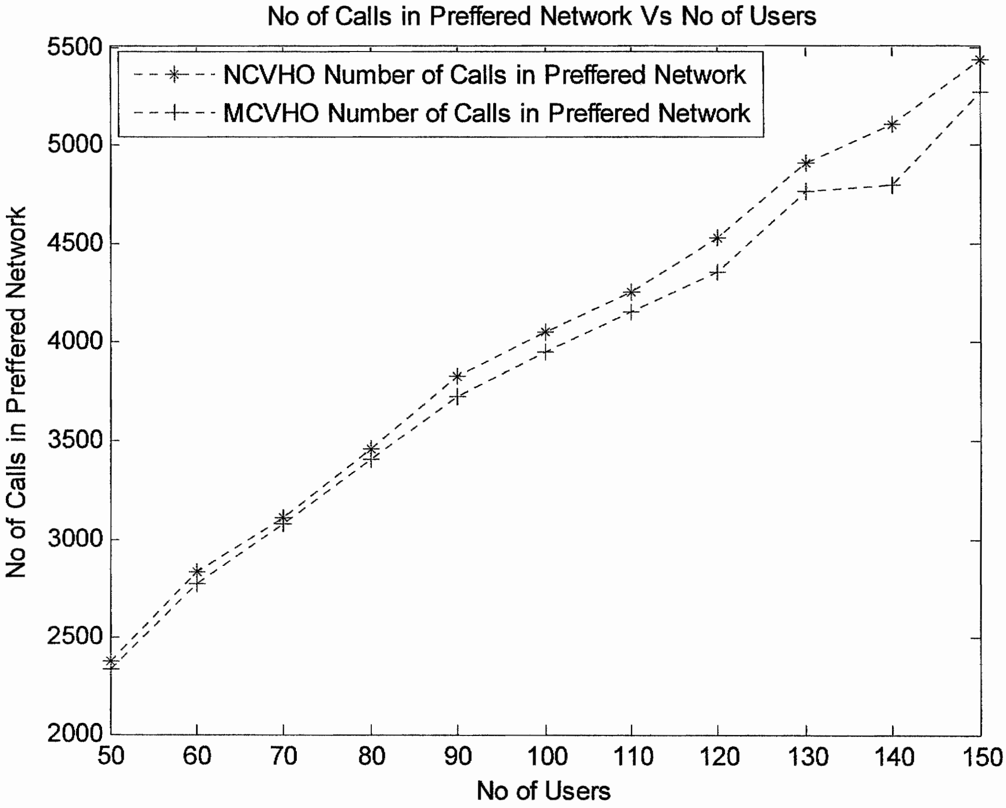


Figure 5.21 Number of calls made in a preferred network Vs Total Number of Users

To analyze user satisfaction, for each service that a user can access, a list of preferred networks for each service is defined on the basis of cost and quality criteria. During the simulation, each time when a vertical handover takes place, the target network and the type of service are registered.

In Figure 5.21, the number of total calls made in a preferred network of users is shown. As the calls hosted by a preferred network increases, users are satisfied to a greater extent in the network. It can be observed from the figure that both the schemes take care of user preferences, and in case of NCVHO, more calls are hosted by a preferred network. As both the schemes are fed with an exactly same traffic pattern, this result indicates that NCVHO performs better than MCVHO in terms of user satisfaction.

5.13 Summary

In this chapter, vertical handover decision-making schemes are discussed. The discussion included the performance analysis of these schemes and the effect of change in parameters, change in weights of parameters, and change of handover control scheme. The analysis shows that change in any of these factors affects the performance of the handover. The analysis of two handover schemes, NCVHO and MCVHO, in terms of the number of total handovers, call dropping, and user satisfaction is also presented. The results show that NCVHO outperforms MCVHO in terms of all these parameters. Therefore, it is concluded that the NCVHO scheme accounts for less number of handovers and less call dropping, which results in better service and increased user satisfaction.

6 Distributed Network Controlled Vertical Handover

6.1 Introduction

The MCVHO and NCVHO schemes are described in chapter 4. The performance analysis of these schemes in terms of handover delay, amount of signaling, usage of radio link, number of handovers, call dropping, and user satisfaction are presented in chapters 4 and 5. Both schemes were identified to have some advantages as well as disadvantages and one performs better than the other under certain situations. NCVHO was found to enhance the performance of a handover by using more information and less signaling on the wireless part of the network. However, the performance of NCVHO has deteriorated as a result of the increased processing delay because of the centralized decision making approach when the number of users is high. In this chapter, the limitations of MCVHO and NCVHO are discussed. Furthermore, enhancements to the NCVHO scheme in order to improve its performance are presented. To achieve this purpose, a new vertical handover scheme known as Distributed Network Controlled Vertical Handover (distNCVHO) is proposed. The performance analysis of distNCVHO in terms of handover delay, number of handovers, and call dropping has also been presented.

6.2 Limitations of MCVHO and NCVHO

6.2.1 Limitations of MCVHO

When compared with NCVHO in terms of handover delay (presented in Chapter 4), MCVHO achieves better delay performance when the number of users is high and when FER is low. The MCVHO scheme allows a mobile terminal to make a handover decision. This results in distributed processing of a handover and minimizes the effect of the increased number of users on handover delay. However, because of the limitation of information about current state of the network available to the MS, this scheme fails to perform better than the NCVHO in terms of the number of handovers and call dropping.

In addition, its performance in terms of handover delay is also worse as compared with NCVHO when the number of users is low or when FER is high.

6.2.2 Limitations of NCVHO

The NCVHO scheme, introduced in Chapter 4, overcomes some of the limitations of MCVHO. This scheme allows the current network of an MS to make a handover decision on behalf of the MS; this reduces signaling on the wireless part of the network. Therefore, the NCVHO scheme has better performance than MCVHO in terms of handover delay when the number of users is low. Furthermore, the NCVHO scheme uses more information to enhance the performance of a handover; this results in better performance in terms of the total number of handovers as well as call dropping rate. However, the centralized approach of handover decision making by the NCVHO leads to increased processing load of the current Point of Attachment (PoA). As a consequence, the performance of this scheme decreases when the number of users is high. This performance degradation could result in NCVHO facing some scalability issues.

To find out a solution to this problem, some important issues with respect to the signaling of NCVHO must be identified. Figure 6.1, which also appears in chapter 4 as Figure 4.6, is repeated here to better explain the key points.

To make enhancements, first, a complete procedure of handover that is involved in processing at current PoA (oPoA) is discussed and the key points related to enhancement are identified.

In the first part of the NCVHO scheme, the MS sends a vertical handover request to oPoA. This request has the information about candidate PoAs, which is available to the MS. After the request has been sent, the information must be processed by the oPoA to shortlist candidate PoAs.

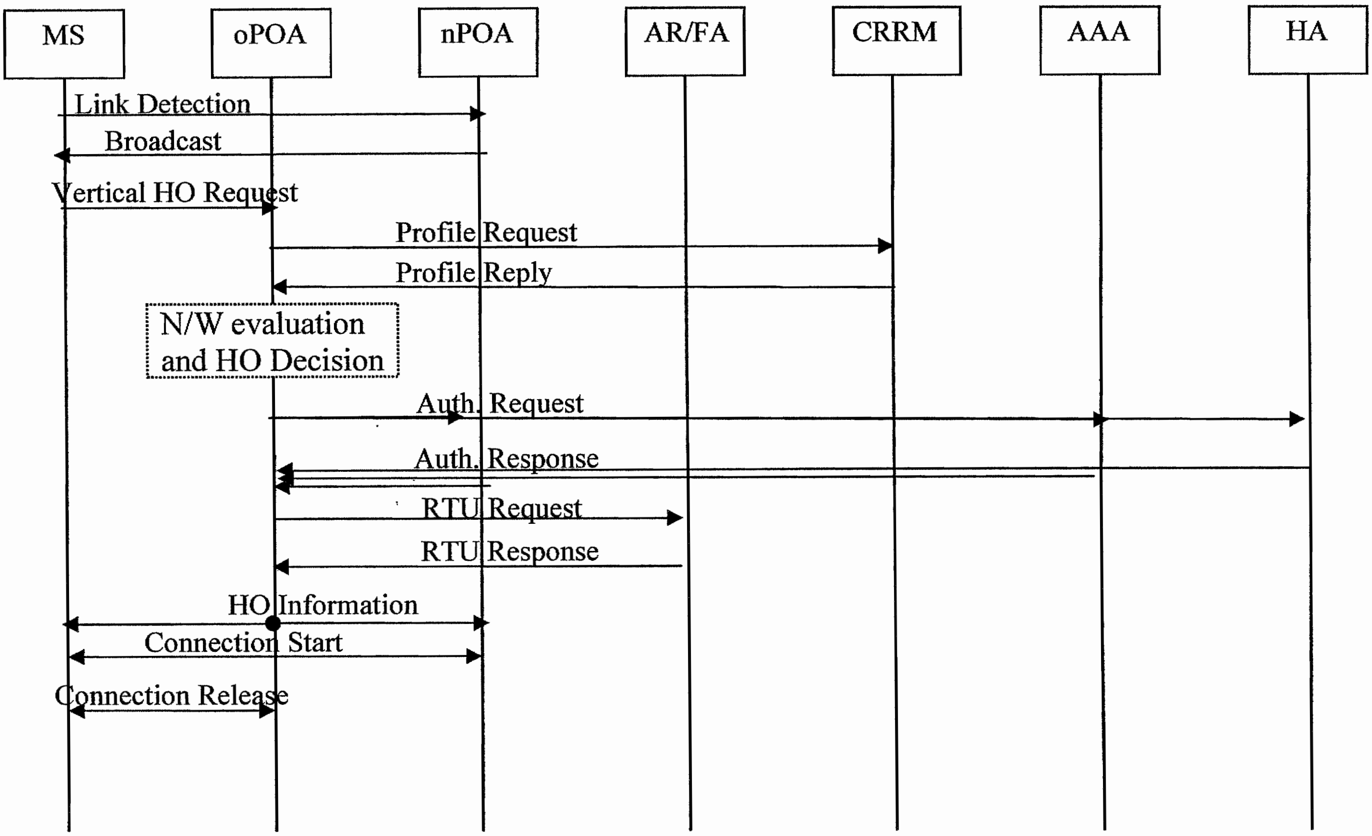


Figure 6.1 Network Controlled Vertical Handover (NCVHO)

To evaluate the candidate PoAs, the oPoA also sends a profile request to the CRRM entity, which sends the profile of candidate PoAs to the oPoA. In UMTS, Node B processes this information also.

Information is thus distributed in the network. The information regarding candidate PoAs is available to the MS and the CRRM entity and is processed together at oPoA. As a result, the processing load at oPoA is increased.

6.3 Resolution

To resolve the issues regarding MCVHO and NCVHO, some enhancement to the schemes is needed. The NCVHO scheme performs better than MCVHO under most situations and responds better to most network conditions. The performance of this handover scheme is limited by the processing load at the current PoA. To enhance the performance of this scheme, either the processing capability of the PoA has to be increased or the processing load of the PoA should be decreased. For achieving this goal, an approach to minimize the processing load of PoA is considered in this chapter. To distribute the processing of information, it is a good idea to process information where it is available. The enhancement of the NCVHO scheme presented here follows the approach of distributed processing of information to reduce the processing load of PoAs. This new scheme is known as distNCVHO because it distributes processing in the basic NCVHO scheme. This scheme is targeted to minimize the handover delay when the number of users is high without sacrificing the performance of NCVHO in other situations. This scheme is described in the following section.

6.4 Distributed Network Controlled Vertical Handover (distNCVHO)

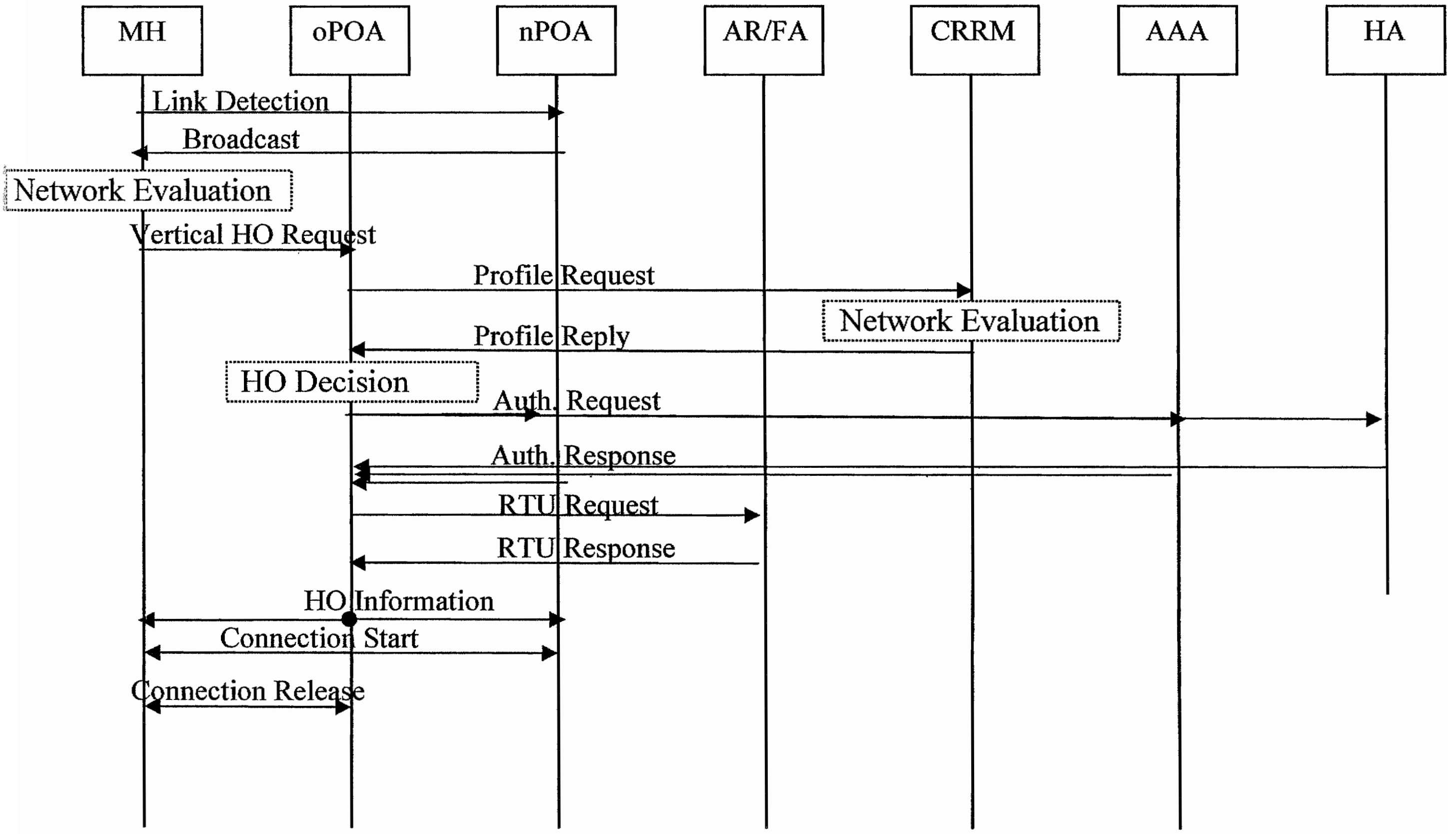


Figure 6.2 Distributed Network Controlled Vertical Handover

The distNCVHO scheme, presented in Figure 6.2, uses the same signaling as NCVHO. As mentioned earlier, this scheme changes the processing of information at different steps of the procedure as explained in the following text.

In the first part of the handover, the MS gathers beacons from other PoAs. These beacons provide information about signal strength and the channels available in the candidate PoAs. In addition to this information, the MS also knows the type of network to which the PoA belongs and it has its own preferences. Combining the information available from the candidate PoA and in itself, the MS creates a sorted list of candidate PoAs, which will be sent to the current PoA along with the vertical handover request.

The current PoA sends a profile request to the associated CRRM entity, which triggers a processing at the CRRM entity. This processing uses the information available to CRRM about candidate PoAs and their networks. The CRRM entity uses all the information it has about the candidate PoA and creates a sorted list of candidate PoAs according to network evaluation. The generated list is in the same format as the one sent by the MS to the current PoA. This list is now sent as a profile reply to the current PoA

After receiving the profile reply from the CRRM, the current PoA has two sorted lists of candidate PoAs, which have to be merged to create a final list. This merging can be either simple by giving greater preference to the list created by the MS or done using a complex policy. This merging is always done in a simple way in distNCVHO. The current PoA then takes the sorted list that is received from the MS and merges it with the list received from the CRRM. Thus, network preferences are also considered and user preferences are also taken into account. Simple merging is easy to perform and does not create significant processing delay.

The rest of the signaling and the procedure of NCVHO are adopted as they are by distNCVHO.

6.5 Performance Analysis

The performance of distNCVHO is analyzed in terms of handover delay, number of total handovers, and call dropping. The analysis uses the similar simulation environment presented in Chapter 4. The signaling delay associated with distNCVHO is given by the following equation. The same notations used in Chapter 4 are used here.

$$T\_distncho = T\_mr + 2*(T\_ra+T\_af) + Tn\_areq + T\_auth + Tn\_arply + Tn\_rturq + Tn\_rturp + T\_mr \quad (6.1)$$

According to Equation 6.1, the distNCVHO scheme uses the same signaling as NCVHO. However, the processing time is reduced significantly, which reduces the overall delay in handover. The performance analysis of distNCVHO in terms of handover delay is done under a similar simulation environment as demonstrated in Chapter 4. Figure 6.3 shows the handover delay of distNCVHO and MCVHO.

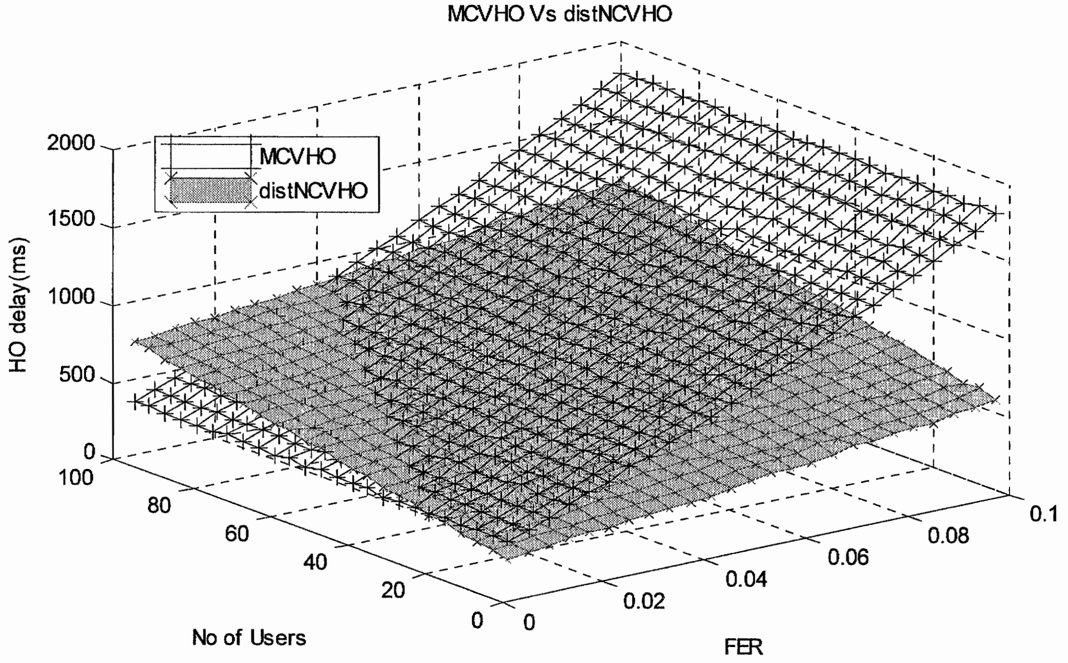


Figure 6.3 Handover Delay of MCVHO Vs distNCVHO

As can be seen in Figure 6.3, the handover delay of distNCVHO is lower than that of MCVHO in most of the conditions. This handover delay is higher than MCVHO only when the number of users is >50 and the FER <0.015. This indicates that the distNCVHO performs better than MCVHO under most of the conditions. The performance of distNCVHO with NCVHO is compared in Figure 6.4, which also shows the handover delay performance of the MCVHO, NCVHO, and distNCVHO schemes.

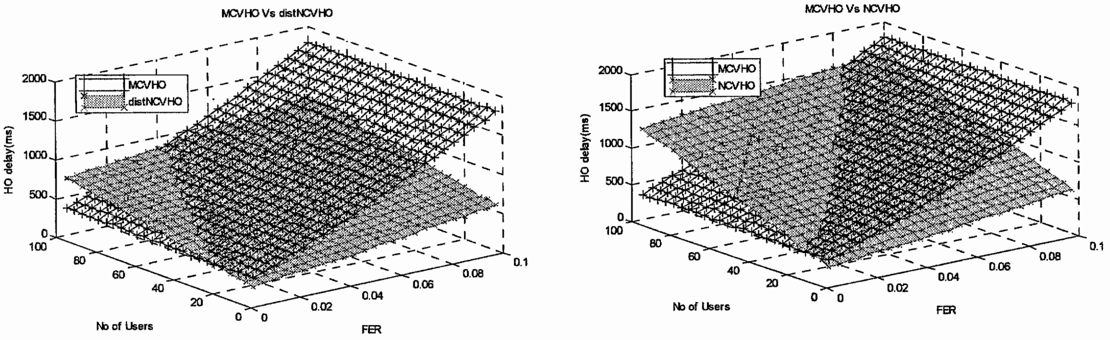


Figure 6.4 Handover Delay distNCVHO Vs MCVHO and NCVHO Vs MCVHO

According to Figure 6.4, the performance of distNCVHO is higher than that of NCVHO under all the conditions. The handover delay associated with distNCVHO is nearly half than that of NCVHO when the number of users is high. This indicates that the distribution of processing has improved the handover delay performance of the scheme significantly. This new scheme allows better utilization of resources in the network by distributing processing between different nodes and achieves better delay performance. The effect of FER seems to be similar on distNCVHO and NCVHO.

This increase in the performance in terms of handover delay can also increase its performance in terms of call dropping because now the network has to serve less load allowing more handovers of users to be served by the network. To analyze the performance of distNCVHO in terms of the number of handovers and call dropping, a setup for performance analysis similar to the one described in Chapter 5 is used. The performance analysis in terms of the number of handovers and call dropping is shown in Figures 6.5 and 6.6, respectively.

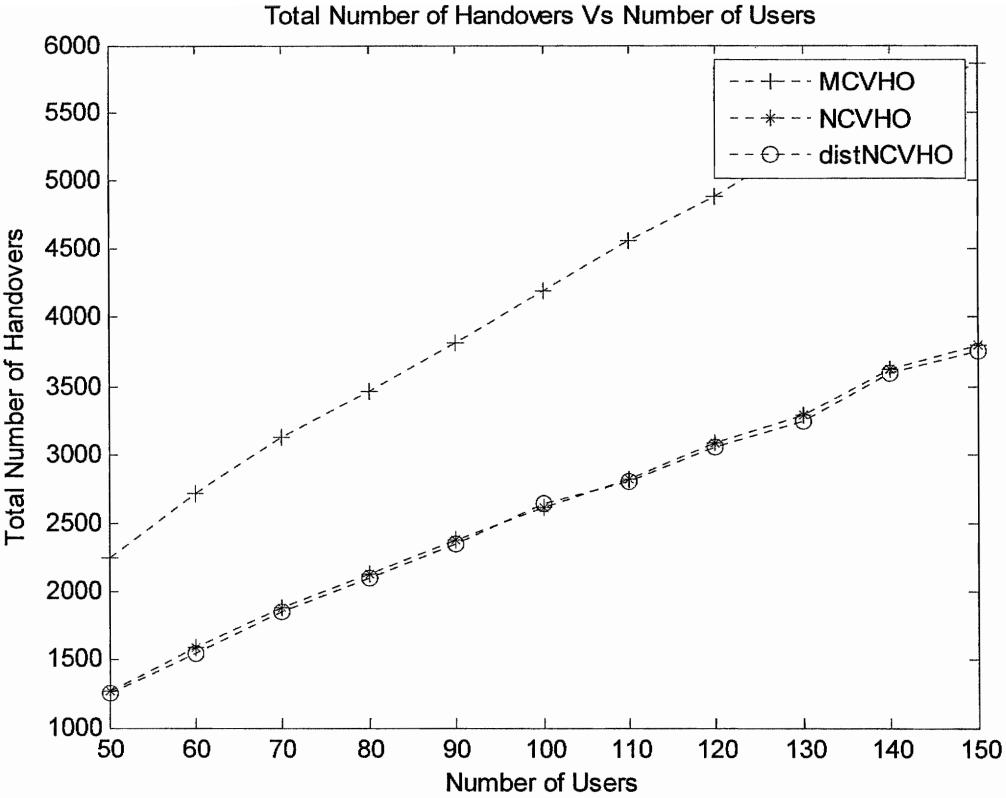


Figure 6.5 Total Number of Handovers Vs Number of Users

The performances of MCVHO, NCVHO and distNCVHO in terms of the total number of handovers during the simulation period are shown in Figure 6.5. The graph shows the total number of handovers during the entire simulation period. The simulation was run 100 times for each number of users with 120 seconds mean call holding time and then the average of the total number of handovers during the entire simulation period was taken and plotted It can be seen from the graph that MCVHO has the highest number of handovers for all number of users. The number of handovers for the NCVHO and distNCVHO schemes is nearly the same for each number of users, which are lower than MCVHO by a large margin. This graph indicates that the distNCVHO achieves almost similar performance as NCVHO in terms of the total number of handover. Therefore, a reduction in handover delay does not adversely affect the performance of distNCVHO in terms of the total number of handovers.

In the next step, the performance of distNCVHO in terms of call dropping is observed. The performances of MCVHO, NCVHO, and distNCVHO in terms of the total number of calls dropped during the entire simulation period for each number of users are shown in Figure 6.6.

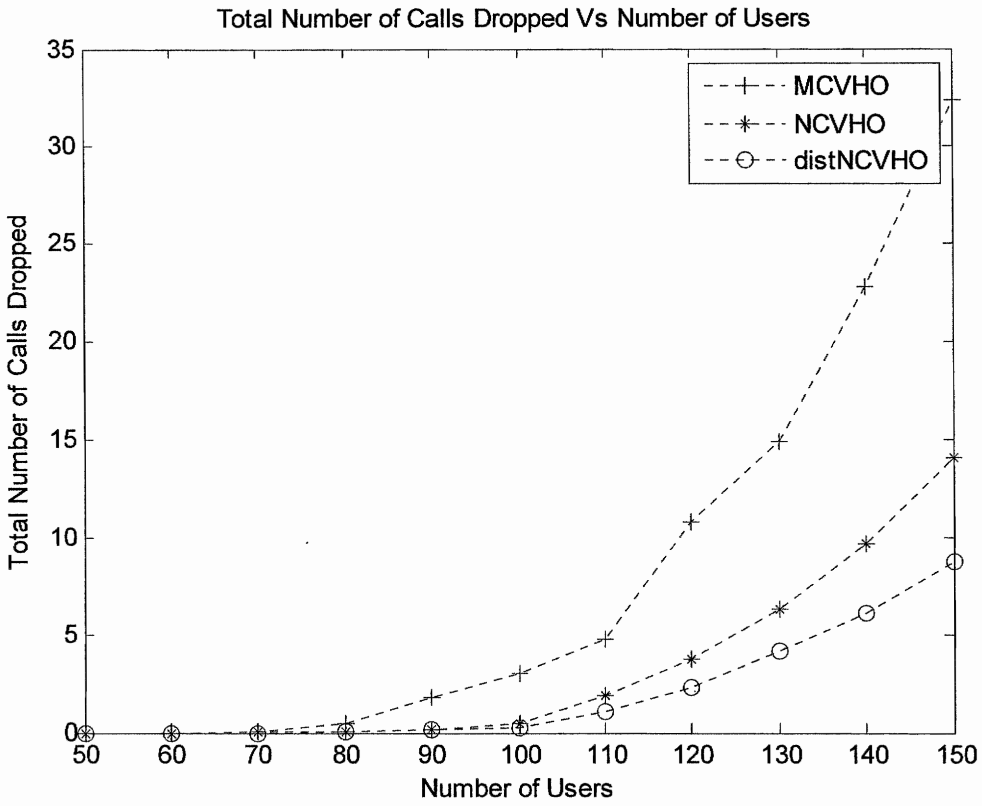


Figure 6.6 Total Number of Calls Dropped Vs Number of Users

Figure 6.6 is also plotted from the same simulation as Figure 6.5. Figure 6.6 indicates the number of calls dropped during the entire period of simulation for each number of users. To check for the consistency of the results, these results are also averaged from 100 rounds of simulation similar to the results shown in Figure 6.5. In Figure 6.6, it can be observed that the MCVHO has the highest number of calls dropped among all the three handover schemes and that distNCVHO has the lowest number of calls dropped. This figure indicates that the performance of distNCVHO is higher than both MCVHO and NCVHO in terms of the number of calls dropped. This result supports the assumption made in the previous section. In the case of distNCVHO, the processing delay of the network during handovers is reduced, which has effectively reduced handover delay. This reduction has also reduced call dropping in the simulation.

6.6 Summary

The limitations of MCVHO and NCVHO are discussed in detail in this chapter. The important issues with respect to MCVHO and NCVHO, which can be addressed, are shown. After identifying that the NCVHO has more benefits than the MCVHO, enhancements to NCVHO are proposed. The major issue of NCVHO of centralized processing is addressed and the distNCVHO scheme that resolves this problem of centralized processing is proposed. From the performance analysis of distNCVHO, it has been shown that the distNCVHO outperforms NCVHO in terms of handover delay under all situations. In addition, it also performs better than the MCVHO in terms of handover delay in most cases. The performance of the distNCVHO scheme is to NCVHO in terms of the total number of handovers, which is better than the performance of the MCVHO scheme. The distNCVHO scheme also has lowest number of dropped calls among all the three schemes.

7 Conclusion and Future Work

7.1 Conclusion

Handover is one of the important radio resource management functions in a cellular network. The primary goal of a handover is to allow users to access services while they keep moving across cells. Horizontal handovers in homogeneous networks are specified for individual networks. Vertical handover for heterogeneous networks are proposed in many literatures. Most literatures propose mobile controlled vertical handovers as a solution to cope with complexities in a heterogeneous network. Some of the limitations of these mobile controlled vertical handovers are identified in this thesis. For instance, the Mobile Station (MS) has a limited battery power in addition to very limited information about current network conditions. Furthermore, in an MCVHO, most of the signaling is performed on the wireless part of the network. Therefore, the performance of an MCVHO is highly affected by the quality of the radio network.

A new Network-Controlled Vertical Handover (NCVHO) is proposed in Chapter 4 with the objective of overcoming the limitations of MCVHO. In this chapter, the performance of an NCVHO is compared with MCVHO and an MIPv6 based vertical handover in terms of handover delay and usage of radio network. The results show that the NCVHO performs better than the MCVHO when the number of users is <40 or when the FER is >0.03. The results presented in this chapter signify that the performance of an MIPv6 based vertical handover is limited by its authentication method. The MCVHO scheme is highly affected by the quality of the radio network and the performance of this scheme degrades when FER increases. Performance of the NCVHO scheme is worse than MCVHO when the number of user increases to >40. These results show that the NCVHO and the MCVHO perform better than each other under different situations in terms of handover delay. Performance of the NCVHO scheme is affected by the increased processing load at oPoA, which increases handover delay when the number of users is high. This signifies that the NCVHO may face scalability problems.

In Chapter 5, performance analysis of the NCVHO and the MCVHO is carried out in terms of the number of total handovers, call dropping, and user satisfaction. A simulation is done for a heterogeneous network that consists of a UMTS network and a WLAN. This chapter includes the study of the effect of parameters used in making a handover decision on the performance of a handover scheme. A cost function based handover decision making scheme is used in this chapter. This scheme takes four parameters into consideration while making a handover decision. The trace of user distribution for a simulation period is used to see how much control a handover scheme has on user distribution between the UMTS network and WLAN. It should be noted that different weights of the parameters with weight tuning achieve optimal performance of a handover scheme. The simulation shows that the NCVHO achieved better control over user distribution than the MCVHO. This was achieved by using extra information about the current condition of the networks from the CRRM entity. The simulation results for the number of handovers and call dropping show that the NCVHO achieved a lower number of handovers and call dropping than MCVHO for any number of users. The results of the number of calls made in a preferred network show that the NCVHO outperforms MCVHO in terms of user satisfaction.

In Chapter 6, the issues related to MCVHO and NCVHO are discussed. It is noticed that the NCVHO performs better than the MCVHO under most situations. However, the problem of increased processing load at oPoA in NCVHO needs to be addressed. In this chapter, a new vertical handover scheme namely the distNCVHO is proposed, which distributed the processing load of oPoA between the CRRM entity and the MS. The simulation results presented in this chapter show that the distNCVHO outperforms NCVHO under all the situations in terms of handover delay. It also achieves better handover delay performance than the MCVHO under most situations. The results for the total number of handovers and call dropping show that the distNCVHO outperforms both NCVHO and MCVHO. The distNCVHO scheme performs better than MCVHO by means of reduced amount of signaling on the wireless part of the network and outperforms NCVHO by distributing processing load between the oPoA, the CRRM entity, and the MS.

7.2 Major Contributions

Three major contributions are made in this thesis in Chapters 4, 5 and 6, respectively.

In Chapter 4, a new vertical handover scheme, namely NCVHO is proposed. The simulation results show that when the number of users is <40 and or when the FER is >0.03 this scheme performs better than MCVHO in terms of handover delay. The scheme also reduces the usage of radio network, which is a valuable resource in a cellular network.

In Chapter 5, the effect of parameters on handover decision making is studied using cost function based handover decision making scheme. The study shows that different weights of parameters with weight tuning achieve the best control over user distribution among different networks. The simulation results of the NCVHO and the MCVHO in terms of the number of handovers, call dropping, and number of calls made in a preferred network show that the NCVHO outperforms MCVHO for all these performance parameters.

In Chapter 6, a novel vertical handover scheme, namely distNCVHO is proposed which outperforms both the MCVHO and the NCVHO in terms of handover delay, number of handovers, and call dropping.

7.3 Future Work

In this thesis, a handover is studied with regard to future heterogeneous networks. As no specification is available for these networks, the information available in literature and widely accepted proposals are used for designing the simulation. However, working groups such as the 3GPP have already started providing specifications for interoperability with other technologies. Specifications such as these can be used to design a detailed network scenario under which a vertical handover can be tested.

In this thesis, all the simulation work is carried out using the MATLAB software. Taking into account the low speed of MATLAB in performing a simulation, many secondary tasks in simulation such as power control, admission control, and congestion control are simplified to achieve results within available period. A detailed simulation scenario with some other faster platform can be used to validate the results obtained in this thesis. In most cases, the simplification is done only where it is unlikely to affect the performance of the proposed vertical handover scheme. Therefore, in most cases, a detailed network scenario will also not affect the results achieved here.

As discussed in Chapter 4, the NCVHO scheme centralizes the processing of handover information at the oPoA. This centralization approach may cause scalability issues for this handover. A solution to this problem is also proposed in Chapter 6 as the distNCVHO scheme. The study of scalability for these two schemes is also an interesting expansion of this study.

8 Reference

[1] N. Tripathi, “Generic Adaptive Handoff Algorithms Using Fuzzy Logic and Neural Networks”, Ph. D. Thesis, Virginia Polytechnic Institute and State University, Virginia, in 1997.

[2] T. S. Rappaport, Wireless Communications. Prentice-Hall Inc., 1996.

[3] M. Mouly and M.-B. Pautet, The GSM System for Mobile Communications. Michel Mouly and Marie-Bernadette Pautet, 1992.

[4] M. Anagnostou and G. C. Manos, “Handover related performance of mobile communication networks," in Proc. 44th IEEE VTC, pp. 111{14, 1994.

[5] A. Kaloxylos, G. Lampropoulos, N. Passas, L. Merakos, (2006) “A flexible handover mechanism for seamless service continuity in heterogeneous environments”, Computer Communications, Vol. 29, pp. 717-729.

[6] W. C. Y. Lee, Mobile Cellular Telecommunications. McGraw Hill, 2 ed., 1995

[7] S Yang, S G Choi, J S Park, H J Kim, (2006) A Handover Framework for Seamless Service Support between Wired and Wireless Networks, Processing of ICACT, Feb. 20-22.

[8] M. Smadi, V. Azhari, T. D. Todd, (2006) A Measurement-Based Study of WLAN to Cellular Handover, IEEE.

[9] V. A. Sousa, R. A. Neto, F. S. Chaves, L. S. Cardoso, F. R. Cavalcanti, (2006) “Access Selection with Anticipatory Vertical Handover for Multi-Access Networks”, IEEE International Symposium on Personal, Indoor and Mobile Radio Communications.

[10] S. Balasubramanian, T. Pfeifer, J. Indulska (2006) Active Node supporting Context-aware Vertical Handover in Pervasive Computing Environment with Redundant Positioning, 1st International Symposium on Wireless Pervasive Computing Vol. 1 , Issue 6, pp. 16-18

[11] H.-H. Choi, O. Song, and D.-H. Cho, (2007) "Seamless handoff scheme based on pre-registration and pre-authentication for UMTS-WLAN interworking," Wireless Personal Communications, Vol. 41, Issue 3, pp. 345–364

[12] H. P. Chang, Q. J. Chen, (2006) An Agile Vertical Handoff Scheme for Heterogeneous Networks, IEEE conference on Pervasive Services, pp. 88-97.

[13] N. Akkari, S. Tohme, M. Doughan, (2007) "Applying Anticipated Vercal Handover in Next Generation Networks" Proceedings of the Fourth European Conference on Universal Multiservice Networks, IEEE.

[14] T Ahmed, K Kyamakya, M Ludwig, (2006) Architecture of a Context-Aware Vertical Handover Decision Model and Its Performance Analysis for GPRS – WiFi Handover, Proceedings of the 11th IEEE Symposium on Computers and Communications.

[15] H. Kwon, K. R. Jung A. Park, J. C. Ryou, (2005) "Consideration of UMTS-WLAN Seamless Handover", Proceedings of the Seventh IEEE International Symposium on Multimedia

[16] W. Zhang, J. Jaehnert, K. Dolzer, (2003) "Design and Evaluation of a Handover Decision Strategy for 4th Generation Mobile Networks", Institute of Communication Networks and Computer Engineering, University of Stuttgart.

[17] T. Ahmed, K. Kyamakya, M. Ludwig, (2006) Design and Implementation of a Context-Aware Decision Algorithm for Heterogeneous Networks, SAC, Dijon, France, April 23-27.

[18] C. M. Chou, C. Y. Huang, (2006) "Dynamic Vertical Handover Control Algorithm For WLAN and UMTS", WCNC'2006, IEEE.

[19] C. Makaya, S. Pierre, (2007) "Efficient Handoff Scheme for Heterogeneous IPv6-based Wireless Networks", WCNC'07.

[20] R. Pintenet, J. Gosteau, T Gizawi, F. Lazarakis, K. Peppas, A. Alexious, (2003) "Evaluation of interoperability mechanism for coexisting HSDPA and WLAN enhanced with MTMR techniques", IEEE.

[21] J. G. Atallah, M. Ismail, (2006) "Future 4G Front-Ends Enabling Smooth Vertical Handovers", IEEE Circuit & Design Magazine, January/February, 2006

[22] D. Axiotis, T. Al-Gizawi, J. Gosteua, F. Lazarakis, K. Peppas, R. Pintenet, A. Moussa C. Vlahodimitropoulos, "FITNESS Simulation Platform Structure and System Level Performance Evaluation", Information Science Technologies, Athens, in 2003.

[23] J. Perez-Romero, O. Sallent, "Loose and Tight Interworking between Vertical and
[24] Horizontal Handovers in Multi-RAT Scenarios", Benalmadena, Spain in 2006

[25] F. Siddiqui, S. Zeadally, "Mobility management across hybrid wireless networks: Trends and challenges" Computer Communications in 2005, Vol. 29, pp.1363-1385

[26] McNair, F. Zhu. "Vertical handoffs in Fourth Generation Multi Network Environments", University of Florida, IEEE Wireless Communications in June 2004

[27] Q. Song, A. Jamalipour, “Network Selection in an Integrated Wireless LAN and UMTS Environment Using Mathematical Modeling and Computing Techniques”, University of Sydney, IEEE Wireless Communications in June 2005.

[28] F. Akyildiz et al. "Mobility Management in Next Generation Wireless System!." Proc, IEEE, in 1999, Vol. 87. No. 8, pp, 1347-84.

[29] W. W. Lu. "Fourth-Generation Mobile Initiatives and Technologies", IEEE Communication Magazine, in May 2002, Vol. 40, No, 3, pp. 10445.

[30] H. Fathi, S. S. Chakraborty, and R. Prasad, "Optimization of MIPv6 based protocol to support VoIP services in wireless heterogeneous networks," *IEEE Transactions on vehicular technology,* vol. 56, p. 11, 2007.

[31] D. Johnson, C. Perkins, and J. Arkko, Mobility Support in IPv6, Jun. 2004. IETF RFC 3775.

[32] R. G. garroppa, S. Giordano, S. Lucetti, G. Risi, L. Tavanti, “Optimization of Vertical Handover Decision Procedure Using an Experimental MIPv6 Testbed”, IEEE, in 2006.

[33] UMTS-WLAN Integrated Network”, 14th IEEE, International Symposium on Personal, Indoor, Mobile Radio Communication Proceedings in 2003.

[34] A. Alexiou, T. Al-Gizawi, D. Axiotis, D. Bourse, S. Buljore, A. Chatzikonstantinou, L. Elicegui, J. Gosteau, R. Karimi, J. Kruys, F. Lazarakis, K. Peppas, R. Pintenet, S. Simoens, “System-Level Simulation Methodology Defined”, Information Science Technologies, Athens in 2002

[35] S. N. P. Van Cauwenberge, “Study of Soft Handover in UMTS”, Ph.D. Thesis, Technical University of Denmark, in July 2003.

[36] A. Alexiou, T. Al-Gizawi, D. Axiotis, J. Gosteau, R. Karimi, F. Lazarakis, A. Moussa, K. Peppas, R. Pintenet, M. Qaddi, K. Vlahodimitropoulos “UMTS/WLAN Interoperability Scenarios and Mechanisms”, Information Science Technologies, Athens, in 2003.

[37] Hou, D. C. O’brien, “Vertical Handover Decision-Making Algorithm Using Fuzzy Logic for the Integrated Radio-and-OW System”, IEEE Transactions on Wireless Communications, in 2006, Vol. 5 No. 1.

[38] Z. Sun, J. Sauvola, “Modeling Techniques for Simulation-Based Performance Evaluation of Mobility Management Schemes”, Finland: University of Oulu, in 2004.

[39] P. K. Kaveh Pahlavan, and Jacques Beneat, "Wideband Radio Propagation Modeling for Indoor Geolocation Applications," IEEE Communications Magazine, 1998.

[40] Laasonen, "Radio Propagation Modeling," in Algorithms for ad hoc networking, 2003.

[41] F. C. Massimo Bernaschi, Giulio Iannello, "Vertical Handoff Performance in Heterogeneous Networks," in International Conference on Parallel Processing Workshops: IEEE, 2004.

[42] F. A. González, J. A. Pérez, and V. H. Zárate “HAMS: Layer 2 Handoff Accurate Measurement Strategy in WLANs 802.11” in IEEE wireless communications, 2003

[43] Wu, K. Sandrasegaran, “A Survey on Common Radio Resource Management”, The 2nd International Conference on Wireless Broadband and Ultra Wideband Communication, 2007.

[44] Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. P. Makela, R. Picha, J. vallstrom, “Handoff in Hybrid Mobile Data Networks”, IEEE Personal Communication, 2000

[45] Stemm, H, Kats, “Vertical handoffs in wireless overlay networks”, Mobile Networks and Applications, in 1998, Vol. 3. pp. 335-350.

[46] J. Perez-Romero, O. Sallent, R. Augusti, P. Karlsson, A. Barbaresi, L. Wang, F. Casadevall, M. Dohler, H. Gonzalez, F. Cabaral-Pinto, “Common Radio Resource Management: Functional Models and Implementation Requirements”. University Politecnica de Catalunya, Spain, 2004.

[47] R. Patachaianand, “Adaptive Power Control for UMTS”, Masters Thesis, University of Technology, Sydney, in March 2007.

[48] P. Vidales, “Seamless Mobility in 4G Systems”, Technical Report, University of Cambridge, in November 2005.

[49] G. P. Pollini, “Trends in handover design," in IEEE Communications Magazine, pp. 82-90, March 1996.

[50] A. J. Ransom, “Handoff considerations in microcellular systems planning," in Proc. PIMRC, pp. 804-808, 1995.

[51] D. Munoz-Rodriguez, J. Moreno-Cadenas, M. Ruiz-Sanchez, and F. Gomez Casaneda, “Neural supported handoff methodology in microcellular systems," in Proc. 42nd IEEE VTC, vol. 1, pp. 431-434, 1992.

[52] S. Chia, “The control of handover initiation in microcells," in Proc. 41st IEEE VTC, pp. 531-536, 1991.

[53] M. Hata and T. Nagatsu, “Mobile location using signal strength measurements in cellular systems," in IEEE Trans. on Vehicular Technology, vol. VT 29, no. 2, pp. 245-252, 1980

[54] G. D. Ott, “Vehicle location in cellular mobile radio systems," in IEEE Transaction on Vehicular Technology, Vol. VT-26, No. 1, pp. 430-436, Feb. 1977

[55] C.N. Chuah, R. D. Yates, and D. Goodman, “Integrated dynamic radio resource management," in Proc. 45th IEEE VTC, pp. 584-588, 1995

[56] M. Etoh, “Next Generation Mobile Systems 3G and Beyond”, John Wiley & Sons, England, in 2005

[57] J. Manner, M. Kojo, “Mobility Related Terminology”, June 2004, IETF RFC 3753