HANDOFF MANAGEMENT IN THE MOBILE WIRELESS INTERNET

[illegible]

[illegible]

HANDOFF MANAGEMENT IN THE MOBILE WIRELESS INTERNET

Bheki Agrippa Cwele

A dissertation submitted in fulfilment of the requirements for the degree of

Master of Science (Computer Science)

Department of Computer Science, Faculty of Science and Agriculture,
University of Zululand

2006

DECLARATION

I, *Bheki Agrippa Cwele*, declare that this dissertation represents the author's work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from published or unpublished work of others has been acknowledged in the text and a list of references is given.

Signature of Student

DEDICATION

I dedicate this dissertation to my late father: Mpiyakhe Cwele (1990), my late mother: Mawo Cwele (1992), and my late grandmother, Cholani Cwele (1993).Your words of encouragement and your best wishes for my education, still sound as if they were said only yesterday.

ACKNOWLEDGEMENTS

First and foremost, I thank God for keeping me alive. Through his mercy, he gives me daily strength. God, you are great, you sent us the *Holy Spirit, Shembe,* to guide and teach us your word.

I would then like to thank the following people for their support throughout my research work:

The head of Department of Computer Science, Prof. M.O Adigun: the role you played in building my carrier is appreciated. Your fatherly advice and guidance is also highly appreciated.

My supervisor: Dr S.S Xulu, thank you for every effort, and patience in guiding me throughout this research work. Your supervision is appreciated.

I am also grateful to my co-supervisor: Mr G.E Ojong. Regardless of the load of work you carried in lecturing, you always had time to suggest direction where I was lost in my research. Your encouraging advice and much patience in editing my work is highly appreciated.

I would also like to thank my research colleagues. The knowledge we shared, and the discussions we conducted during the past two years of research, contributed a lot.

Finally and most importantly, I would also like to thank my sponsor, Telkom, for their funding of this research project.

ABSTRACT

The large number of nodes in wireless networks and the fact that mobility is an important feature of a wireless system has made it a great challenge to sustain ongoing connections during handoff events. Real-time applications (e.g., Voice-Over-IP (VoIP), Video-On-demand, and streaming) require little or no service disruptions during handoff because of their nature of being mostly delay-sensitive applications. The network must, therefore, offer minimal handoff delays, minimal latency, low call dropping probability; and should also fully be supportive of speed variations in mobile nodes during network layer handoff.

In our research we developed and simulated a fast network layer (subnet-to-subnet) handoff management scheme, known as *Sector Aware handoff* (SAH) strategy. SAH uses cell sectors marked into regions. The sectors have non-handoff region, preparation region and handoff execution region. The SAH scheme uses mobile node's real-time mobility parameters over cell sector regions, to provide accurate and early predictions of Network Layer (L3) handovers – thus initiating L3 handoff preparation appropriately early. SAH is proposed to support primarily, real-time applications' handoff requirements and also fast moving mobile nodes (that is, supports speed variation in mobile nodes).

As a means of evaluating the performance of the proposed scheme, a simulation was conducted. The experimental simulation results show that SAH performs better in terms of providing low call dropping probability and maximum system utilisation, when compared to the Fast Hierarchical Mobile IPv6 (FHMIPv6) handoff protocol.

Table of Contents

List of Figures

List of Tables

List of Equations

Chapter One

INTRODUCTION

1.1 Overview

The mobility of users of a wireless network is the most important distinguishing feature of a wireless network from its wired counterpart. The wireless network environment has led to a proliferation of wireless mobile devices (laptops, PDAs, 3G phones, etc). Mobile device users today enjoy the freedom of ubiquitous connection, either to stationary hosts or to other mobile hosts, while they are on the move. This has led to a tremendous growth in the usage of mobile devices for running applications over the Internet. Applications that are run by mobile nodes while roaming across wireless cells can be of two main forms, namely: real-time applications and non-real time applications. Real-time applications (e.g., Voice-Over IP (VoIP), Video-On-Demand, Internet telephony) are delay-sensitive. These applications are sensitive to packet delays, packet loss, and high buffering. They can only tolerate very little service disruption (e.g. during handover); otherwise their Quality of Service (QoS) requirements can be violated. Non-real-time applications (e.g. HTTP files, short-message-service (sms), e-mail) are delay-tolerant. In these applications packets can be delayed, or buffered and then later forwarded to the receiving node.

Real-time applications are becoming increasingly popular. There are so many reasons behind this. Here we provide some examples of the usage of real-time applications: A person is travelling and needs some entertainment with his IP-compatible mobile device. He connects to the Internet, downloads and watches streaming video. Also one may be rushing to work in the morning; just then an urgent situation arises that requires her to deposit some cash to a particular account. One would simply use one's mobile device to log onto the Internet banking service, and transfer the money. This prevents one from being late at work and at the same time helps one attend to one's emergency.

However, there are still some challenges that are being faced by wireless networks, which would hinder the networks from rendering of good service to roaming mobile nodes and their sophisticated applications; especially real-time applications. On the one hand, wireless networks function under scarce resources, such as bandwidth, and lower transmission speed in the radio links. On the other hand, real-time applications carried by mobile nodes require some QoS guarantees for them to operate effectively. In the next section we introduce the wireless Internet before defining our research problem in a later section.

1.2 The Wireless Internet

The wireless network environment is made up of cells. (See Figure 1.1 for a typical representation of a wireless access network). The access network is formed by a group of cells that are adjacent to one another. The base station of these cells are linked to a common access router (AR1 and AR2 in Figure 1.1), acting as the gateway to the core network (that is, the global Internet). The access network components are: (i) Base Station (BS) - a transceiver node that emits radio signals to mobile nodes residing in its service area. (ii) Mobile Node (MN) - a station or node that traverses cells and is able to send and receive data. (iii) Cell area – a service area that Base Stations' broadcasted radio waves can diffuse throughout. (iv) Access router (AR) – an Internet Protocol (IP) enabled node that links a group of base stations to the external network. A group of cells, in combination with the access router that links them, can be referred to as a subnet. In Figure 1.1, two access routers AR1 and AR2 each control a subnet.

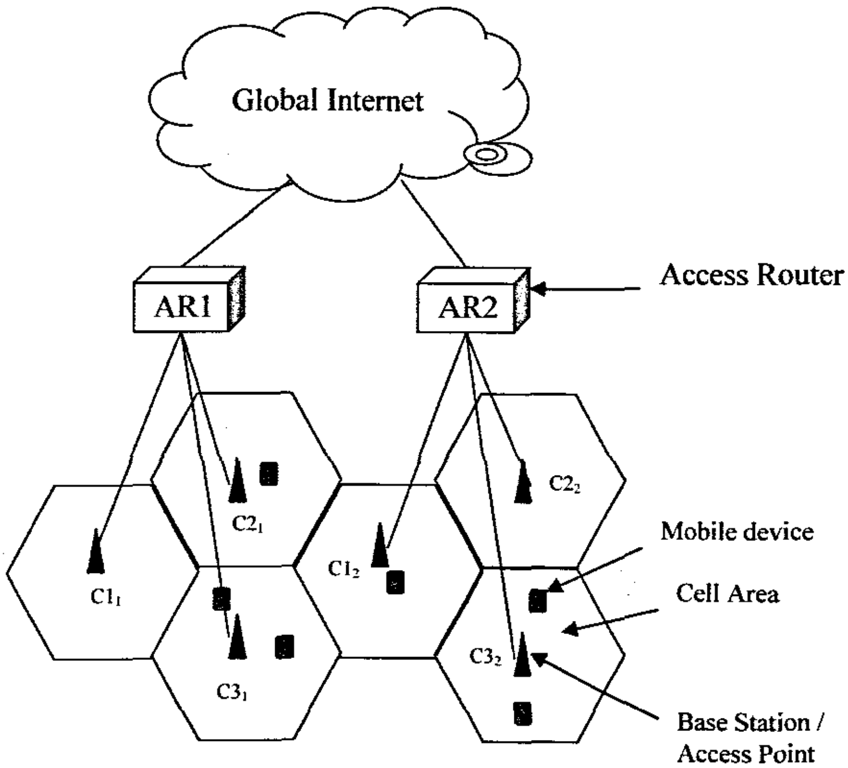

Figure 1.1 Typical representation of Wireless Access Network Environment

In a wireless environment, users can communicate while on the move. Adjacent cells operate at distinct frequency levels with the aim of preventing interferences. This operational characteristic forces moving mobile nodes that are engaged in open connections to switch their active connections each time they move to another cell.

Mobile device users are able to access Internet applications and services through their mobile devices even when they are in motion. They can engage in real-time active communications (conversations - Voice-over-IP, streaming videos, etc). As these mobile devices move, their received signal strength becomes weaker and weaker as they approach cell boundaries. They eventually come to a point where they have to cross these cells boundaries, thus having to change their point of attachment from one base station (BS) or access point (AP) to another. This requires a procedure to transfer the ongoing communication from the old cell's base station to the new cell's base station. This process of transferring an active connection from one point of attachment to another is called *handoff*. In access networks, generally there are two instances of handoffs. There is link layer (L2) handoff and network layer (L3) handoff. In link layer handoff, the mobile node changes only the access point of attachment, i.e. moves between cells that are under the same access router. For example, in Figure 1.1, the mobile node could move from cell $C1_1$ to cell $C3_1$ (which are both under router AR1). In network layer handoff the Mobile Node changes both access point and access router (that is, in link layer, and IP layer). For example, in Figure 1.1 the mobile nodes can move from cell $C3_1$ under AR1 to cell $C1_2$ under AR2.

In general, a good handoff strategy with respect to real-time services, should have the following basic characteristics (Shyy, 2004): *(1) minimal packet loss* – Packet loss is the number of mis-tunnelled packets that eventually do not reach MN. Previously proposed handoff strategies avoid packet loss by making use of buffering at previous access router or new access router. Unfortunately, real-time services can tolerate very little buffering; approximately 5 % for voice. In terms of packet loss for data transmissions, can tolerate as low as 1 % of packet loss (Shyy, 2004). *(2) minimal handoff delay* – Handoff delay is the time taken from the moment a handoff is supposed to occur to the time it actually occurs. This could be introduced either by a delayed router advertisement from the new access router, or by intentional delay techniques (e.g. threshold) as a means to avoid pre-mature handoffs and ping-pong effects (Yang, 2005). These can affect high speed travelling mobile nodes; hence calls may be subjected to high call dropping probability, and degraded signal quality. (3) *minimal latency* – Latency is the time the mobile node last received a packet from the old link to the time it received the first packet from the new link. The above characteristics are very important for a good handoff management strategy to support real-time and non-real-time traffic.

The main focus of this research is to propose a fast, advanced proactive handoff strategy for network layer (L3) micro-mobility. By proactive we mean that we use some measurements to predict and prepare for the impending L3 handoff before its occurrence. In our advanced proactive approach, we initiate handoff preparation early enough before the availability of data link layer (L2) triggers.

In the literature, there are a number of handoff schemes which have been proposed to improve handoff management strategies in order to provide faster handoffs and minimal packet losses, e.g. (Dommety et al, 2002; Tan et al, 1999; Jung et al, 2005 [a], 2005 [b]). These strategies try to solve handoff problems in order to support real-time applications in wireless networks during handoff. Unfortunately current handoff strategies (with regard to network-layer handovers) do not meet delay requirements demanded by real-time applications. Emerging real-time applications evolving from new mobile technology advances, demand stringent QoS maintenance, especial during handovers. Real-time services often carry time-sensitive data that are critical and urgent, for example in monitoring air traffic, remote banking, etc. This requires time-sensitive handoff strategies.

This research work proposes a Sector Aware Handoff (SAH) strategy for L3 handovers. Its approach is similar to Fast Handovers (in being proactive), but uses cell sectors with regions for detection and initiation of L3 handovers. The cell sectors are marked into non-handoff region, handoff preparation and execution regions. L3 handoff prediction occurs anywhere within the cell sector, as opposed to relying only on L2 triggers as the case in other proposed Fast Handovers. SAH is primarily meant to support real-time applications, and their varying QoS requirements during handovers, especially when transmitted by high speed mobile nodes.

1.3 Statement of the Problem

Previously proposed handoff management schemes do not meet network layer (IP layer) minimal handoff delay requirements, required by real-time applications. Furthermore, they do not properly support speed variation of mobile nodes (MNs) traversing across IP subnets. These schemes provide good support to non-real-time applications where service constraints are not very tight. The network-layer handoffs of current proposed schemes still have some problems, which make them less efficient in terms of supporting real-time applications and high speed mobile nodes. Some of the challenges include: (i) The L3 handoff being usually based on the availability of L2 triggers (such as fast handovers in mobile IP). The network layer handoff has no prior knowledge until L2 reveals some indications (e.g., fire triggers) that a mobile node is about to move to the new cell. The previous schemes that are proactive in nature, such as (Lin et al, 2006), and (Saraswady and Shanmugavel 2004), wait for the availability of these triggers (L2 triggers) before initiating L3 handoff. This method is not reliable enough to MNs travelling at high speeds, and also can be less efficient in a situation where there is less overlap in regions between cells. (ii) They rely on signal strengths, as a key determinant for initiating handovers. Real-time services do not merely depend on signal availability but the quality of the available signal. The signal quality depends on signal-to-noise ratio (SNR). The signal strength can be good but if the SNR is low, the signal quality will be low. These characteristics can adversely contribute to handoff delays, and signal-quality degradation. MNs travelling at high speeds can easily lose their previous attachment before handoff is initiated.

This research work proposes an optimal handoff scheme for network-layer handoff to effectively support real-time applications and high speed travelling nodes. Our scheme makes use of mobile node's movement parameters (such as: speed, location, direction, and position) in calibrated cell sectors.

1.4 Research Motivation

The world today is technology driven. M-Commerce is an example that shows how vital it is to have an always-on connection irrespective of the geographical location. Most high ranking businesses maintain their customer relationship through the application of real-time services. Banks have begun to offer their customers Internet banking options (FNB, 2006). In this service, customers can conduct their banking business via their mobile devices (cell phones, laptops, PDA's, etc.) at anytime convenient to them even when mobile. In order for such M-Commerce applications to function properly, there must be a high QoS from the wireless network environment to support real-time applications.

The University of Zululand Computer Science Departmental research also focuses on Quality of Service (QoS) provisioning in the Internet. A good handoff management strategy is one aspect of ensuring high QoS in the wireless Internet. The use of L2 information dependency characteristics for network-layer (L3) handoffs by the previously proposed fast handover schemes, makes them less reliable with regard to the support of real-time applications and fast moving mobile devices. This has led to a motivation to develop a fast and efficient handoff management strategy for L3 handoffs that uses mobile node's movement parameters in real-time, over marked sectors in a cell. A novel L3 handoff strategy that will cater for both real-time and non-real-time applications and also for high speed travelling mobile devices, is necessary.

1.5 Research Goal and Objectives

1.5.1 Research Goal

The goal of this research is to propose a network-layer (Inter-Subnet) handoff management strategy to effectively support real-time multimedia applications and fast moving mobile nodes in the mobile wireless Internet.

1.5.2 Research Objectives

The following objectives formulate the steps to achieve our goal. They are to:

a) identify relevant design characteristics for the proposed network-layer handoff management scheme;

b) formulate a model and define a handoff procedure corresponding to the proposed model and

c) simulate and evaluate the performance of the model through experiments and results interpretation.

1.6 Organisation of the Dissertation

The rest of the dissertation is organised as follows:

Chapter 2 gives the background of the study. It discusses the background concepts of handoff in the Wireless Internet. It provides discussions of different classifications and levels of handoff. Finally, the chapter concludes by outlining the characteristics of a fast handoff.

Chapter 3 reviews previously proposed L3 fast handoff management schemes. The fast handoff schemes are reviewed according to certain characteristics based on how they provide fast handoffs. These schemes are categorised, according to the way their handovers are managed. The chapter also outlines some design challenges in Network Layer (L3) handoff strategies. Generic handoff requirements are also introduced. The chapter concludes by summarising the limitations of previously proposed L3 schemes.

Chapter 4 provides the architectural design of the proposed L3 handoff model. It discusses the design goals and principles of the proposed model. It also discusses the handoff procedure relative to the proposed model, dwelling on the two main phases, i.e. handoff preparation and handoff execution. It further details the activities for each phase. Finally, the chapter concludes by providing handoff algorithms pertaining to the procedure for carrying out the handoff requests.

Chapter 5 gives the details of the simulation implementation of the proposed handoff model. It discusses the simulation environment, and the simulated network topology. It also introduces the performance metrics and simulation parameters. The chapter concludes with discussions of simulation results and their analysis.

Finally, Chapter 6 provides a conclusion of the study. It gives a summary of the achievements obtained in the research study. Some short-comings related to the research study are also outlined. The chapter concludes by suggesting some views and directions of the future work.

Chapter Two

BACKGROUND

2.1 Introduction

The wide spread usage of mobile communication devices with stringent applications' service requirements in the Wireless Internet is a great challenge for Quality of Service (QoS) provisioning in the Wireless Internet. Handoff management is one of the most important QoS issues aimed at offering seamless mobility in the wireless network environment. This chapter gives a background of the research work. We discuss handoff in the Wireless Internet. Basic concepts and terminology of handoff, handoff phases, handoff classifications, and characteristics of a fast handoff are presented. The main focus of the discussion is on Network Layer, L3 (subnet-subnet) handoff but some basics of the lower Layer, L2 (link Layer) handoff are also explained to aid a better understanding of L3 layer handoff.

2.2 Handoff in Wireless Network

Basically handoff, can be defined as a process of switching an active communicating mobile node's connection from one transceiver node (BS) to another, as the mobile node moves out-of-range of an old transceiver to the service area of a new transceiver (see figure 2.1).

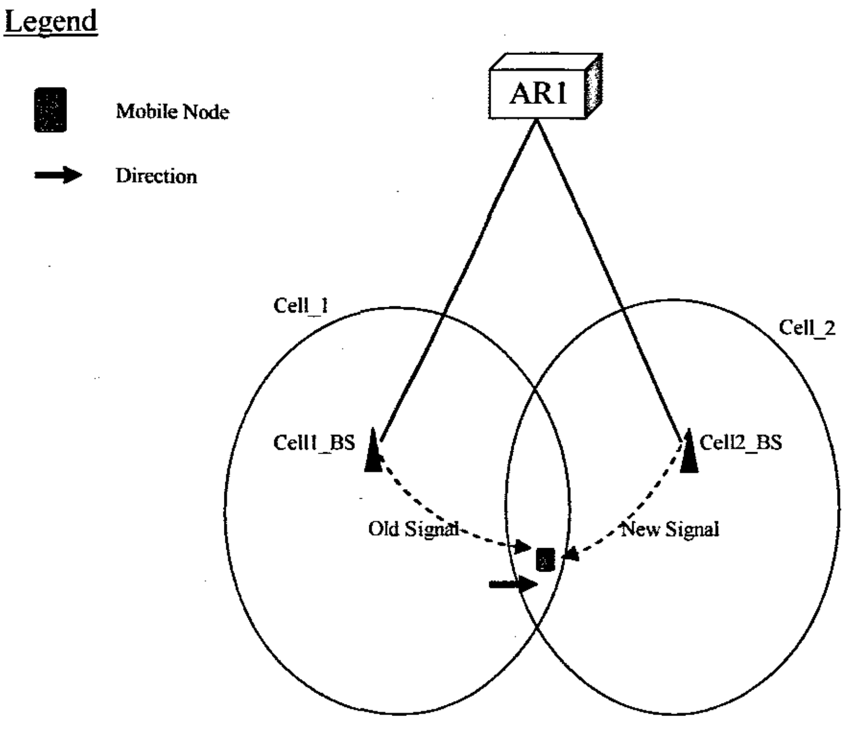

Figure 2.1 Handoff Scenario

In the above diagram (Figure 2.1), a mobile node moves from cell_1 to cell_2. As it approaches cell_1's boundary, the received signal (signal from cell1_BS) becomes weaker and weaker. As it enters the coverage area of cell_2 the new signal from cell2_BS becomes stronger. When the new signal from cell2_BS becomes stronger than the one from cell1_BS, the mobile node will then need to switch to cell2_BS.

There are negative effects that could be introduced during handoff, such as: traffic delays, service disruptions, and call dropping. Handoff management is one of the QoS issues that has attracted the attention of many researchers. This interest in handoff is fuelled by a need to minimise service disruptions, especially for real-time applications as these are the most delay-sensitive applications. The ultimate goal is to achieve fast handoffs with minimal latency, and low call dropping probability, when normal to high speed mobile nodes switch from one point of attachment to another.

Handoff has three basic phases: Handoff initiation, Handoff decision, and Handoff execution.

(a). Handoff Initiation

Handoff can be initiated from various effects, such as signal strength, signal quality, distance from the transmitter, etc. It can also depend on the nature of the application running at the mobile node; for example, real-time applications can rely more on signal quality (i.e. level of signal-to-noise ratio) than just on the signal strength.

(b). Handoff Decision

The handoff decision involves the process of granting or denying the handoff. A good handoff decision is important so as to control false or premature handoffs. False handoffs can occur as a result of fading effects, corner effects, etc. Denying a handoff involves measures such as suspending a requested handoff until it is mature. A threshold value is usually used as the yard stick, where the due handoff can be suppressed by a certain value before it is allowed.

(c). Handoff Execution

This phase comes as the final stage when all the conditions are satisfied. Activities in this phase include resource acquisition, authentications, registrations and switching. The MN is switched from old AP and assigned to the new targeted AP. After the completion of this phase, the handoff can be considered as finished. The status about handoff, whether successful or failed, is also announced in this phase.

In the next section, we discuss the classification of handoffs. Handoffs can be classified in various ways, corresponding to their occurrence. Section 2.2.2 and section 2.2.3 provide detailed discussions on link layer and network layer handoffs respectively.

2.2.1 Classifications of handoffs

Handoff can be classified in various ways depending on: *how it occurs*, *what decides it* (i.e., what network element controls it), and *where it occurs* (i.e., at which layer in the network it occurs).

Firstly, using handoff occurrence, handoff can be classified as either hard or soft handoff. Hard handoff is a "break-before-make" connection type of handoff, while soft handoff is a "make-before-break" connection type of handoff. In hard handoff, the BS or MN stops the connection and starts scanning for a new connection from another BS in the vicinity. When the new connection is found, MN registers with that particular BS. This kind of handoff is typically employed in Frequency Division Multiple Access (FDMA) and Time Division Multiple Access (TDMA) systems. In soft handoff, MN transmits using both links (old and new) until the new link is fully established. Then the old link is terminated. This kind of handoff occurs in Code Division Multiple Access (CDMA) systems.

Secondly, handoff is classified according to which network element controls it. In this light, handoff could be classified as network-controlled, mobile-controlled, or mobile assisted. Network-controlled handoff occurs where the network is responsible for a MN's signal measurements and thus can decide when to initiate handoff. Mobile-controlled handoff occurs where the MN monitors the signal, and upon perceiving degradation of received signal, it requests handoff to another base station with a stronger signal. Mobile assisted handoff occurs where the MN makes measurements and reports to the network (i.e. the BS in-charge), and the network makes the decision. Each of these decisions has its own advantages and disadvantages as far as MN and network capabilities are concerned.

Thirdly, yet another way of classifying handoffs is by the level (layer) at which they can occur in the network (Lopez et al, 2001). It can be categorised as intra-AR (cell-to-cell), inter-AR (subnet-to-subnet), and inter-domain (domain-to-domain). Cell-to-cell handoff involves switching only at the link layer. Subnet-to-Subnet handoff involves switching at both the link layer and the IP layer, in the micro-mobility level (i.e. intra-domain movements). Domain-to-domain handoff involves switching at the IP layer, in the macro-mobility level (i.e. Inter-domain movements). A domain is a group of subnets under the same administration. Figure 2.2 is a graphical representation of handoff levels in a hierarchical structure.

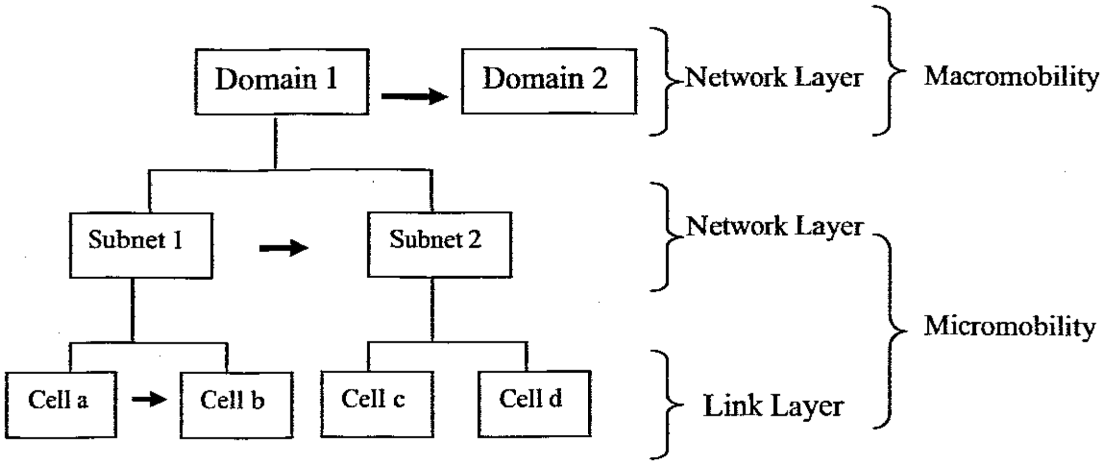

Figure 2.2 Graphical representation of handoff levels

In figure 2.2, a subnet is constituted by a group of cells; for example Cell a, and Cell b belong to subnet 1. Cell grouping can range from 3, 7, 9, 12, 19, or 21, but for simplicity our diagram indicates only 2 cells per subnet. A domain is formed by grouping, for example, subnet-1 and subnet-2. In the next two sections, we shall describe in detail the layered handoffs since they form the backbone of our research.

2.2.2 Link Layer (L2) Handoff

This handoff occurs at the link layer level. The MN merely changes base station or access point of attachment. The base stations or access points belong to the same subnet, i.e. the packets have the same subnet address prefixes. The general determinant of L2 handoff is the signal strength. L2 handoff is usually initiated when the signal strength from the current serving base station falls below the one from the next base station. In Fast Handovers (FMIPv6) (Dommety et al, 2002), L2 handoff information is used by the Network Layer to anticipate its handoff; i.e. in the case where the base stations belong to different subnets. Figure 2.3 illustrates a link layer handoff.

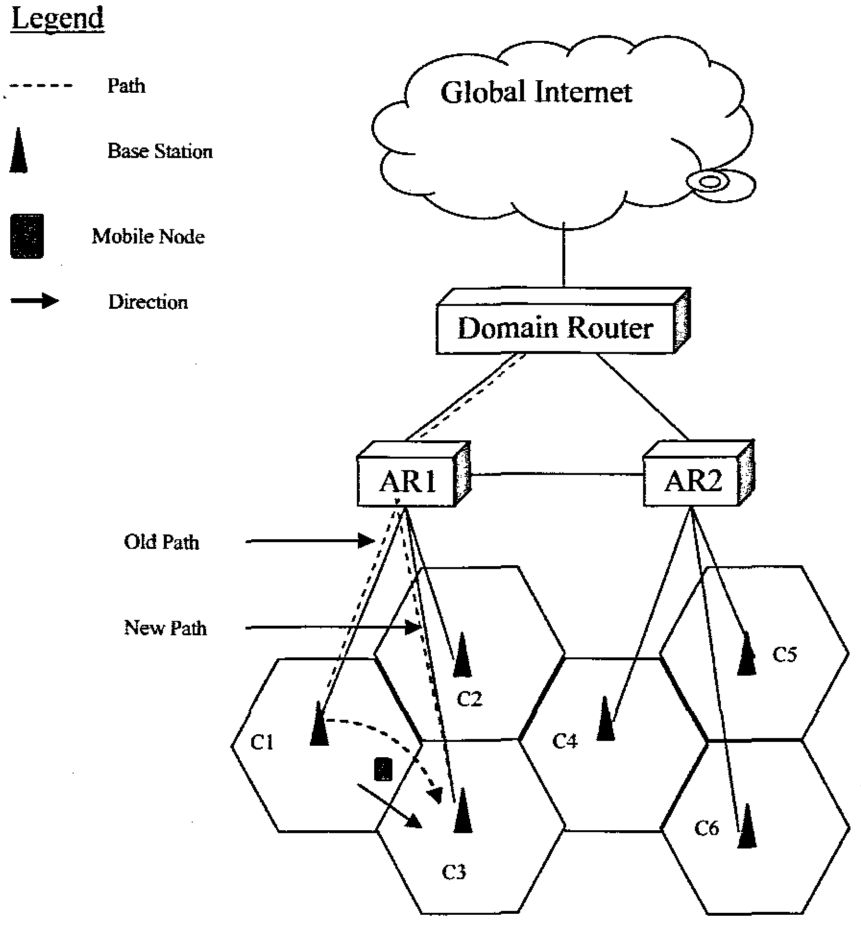

Figure 2.3 Link layer handoff scenario

In the above diagram (Figure 2.3), the access network is composed of two subnets. Each subnet has three cells. Subnet 1 consists of C1, C2, and C3, linked by access router 1 (AR1). Subnet 2 consists of C4, C5, and C6, linked by access router 2 (AR2). Access routers in turn are linked by the domain router. The Domain Router acts as the gateway to the global Internet. The mobile node performs link layer handoff when it moves from cell C1 to cell C3. Both of these cells belong to the same subnet. It only switches between base station or access point. The access router redirects traffic meant for MN to the new cell's BS (C3).

2.2.3 Network Layer (L3) Handoff

The L3 handoff occurs at the network layer level. In this case, the MN moves from a cell of one subnet to another cell belonging to another subnet. The MN changes the base station or access point and also the Subnet IP address. Network layer handoff can also be at the domain level. In this case, the mobile node moves from one domain to another. The handoff at this level usually occurs for macro-mobility and is seldom occurring (since a domain has a large span area). It is normally handled by Mobile IP (Perkins, 1996). The main focus of this study is on the subnet-to-subnet level (inter-subnet handoff). The inter-subnet handoff falls under micro-mobility and is exposed to fairly frequent occurrence of handoffs; as far as high speed nodes and reduced cell sizes are concerned. This handoff has more activities on its occurrence. These include: resource negotiations, transfer of security context, creation of new care-of-addresses, and both Link layer and IP layer switching. (Pack et al, 2002), (Pack et al, 2005), (Shim et al, 2002), (Chou et al, 2005). Therefore it can be deduced that there must be enough time allocated, from its initiation to its execution. Figure 2.4 illustrates the Network Layer (subnet-to-subnet) handoff.

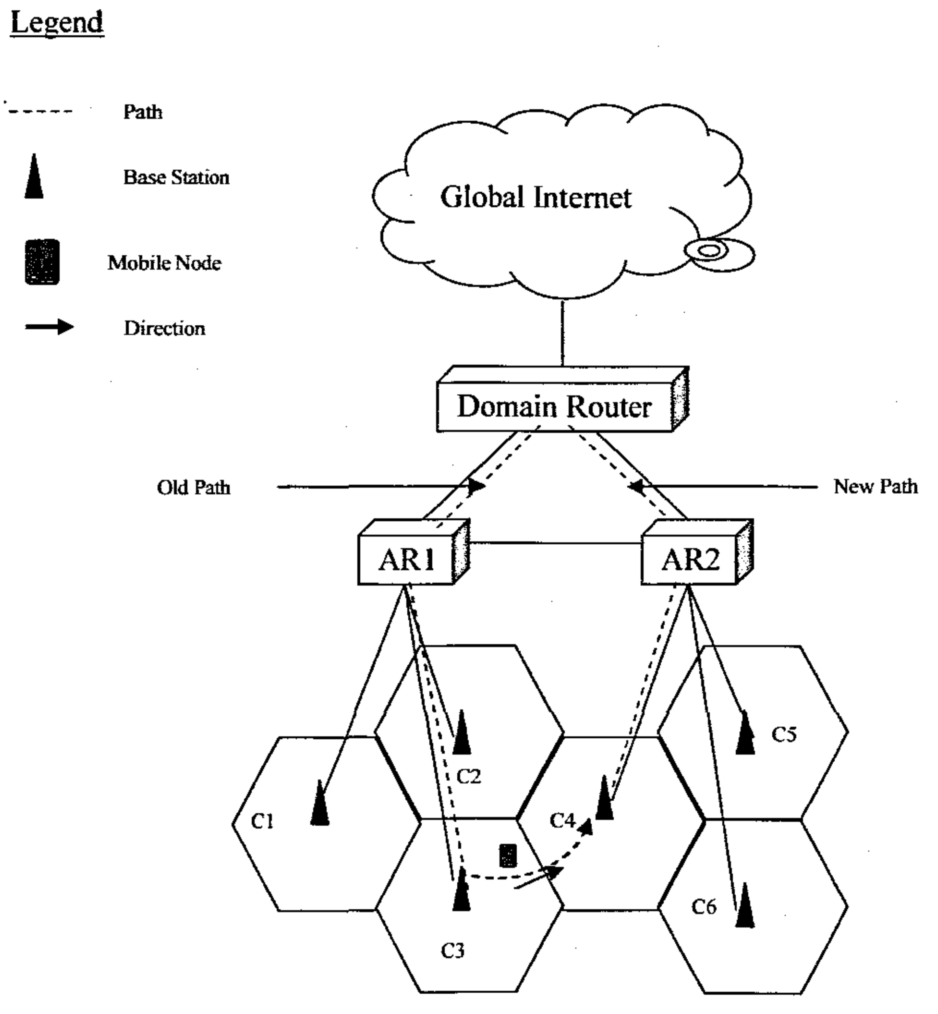

Figure 2.4 Network Layer (subnet-to-subnet) handoff scenario

In the type of handoff shown in figure 2.4, the mobile node moves from cell C3 under the control of AR1 to cell C4 under the control of AR2. Therefore it switches both its link layer and IP layer addresses.

Network layer handoff can follow two approaches: *proactive approach* or *reactive approach*. It also has two registration procedures: *pre-registration procedure* or *post-registration procedure*.

In the proactive approach (Dommety et al, 2002), the network uses link layer triggers to determine imminent handover. This enables the network layer handoff to be initiated before MN loses contact with the old access router (AR). In the reactive approach, the MN waits for router advertisement from the AR. If the new advertisement is different from the previous router advertisement from old AR, then the MN will assume that it has moved to a new access router; hence, initiates handoff or notifies the responsible element to initiate its handoff.

In pre-registration procedure, the mobile node's registration to a new AR is done in advance before MN starts receiving through the new access router. While in post-registration, the mobile node is registered to a new AR after it has finished L2 handover and is under service area of the new AR.

2.3 Characteristics of a fast handoff

Providing fast handoff is a combination of several characteristics. In chapter 3, L3 handoff strategies are reviewed in accordance with the following characteristics: *(i) L3 handoff detection and anticipation procedure, (ii) Handoff resource acquisition procedure.* These characteristics can be described as follows:

(a) L3 handoff detection and anticipation procedure - This involves measures such as *how, where, and when* is L3 handoff decided. The success of any handoff strategy is predominantly based on its handoff detection mechanism. Handoff detection can be *history based, L2 triggers dependent, or location aware.*

(b) *Handoff resource acquisition procedure* - These are the ways in which handoff resources are negotiated at the target cell. They can be reserved at a certain point when L3 handoff has been anticipated, as *in advance reservations* or they can be *immediate* reservations (i.e. request and use now). An another method is *Neighbour casting and Context transfer.* In this procedure, adjacent cells form a group. Once a MN registers with one cell in the group, its context is automatically broadcasted to other cells. Thus the other cells also receive and buffer the most recent packets for MN.

These characteristics contribute in achieving fast L3 handovers, depending on how they are managed. A fast handoff can be characterised by providing *minimal latency, low or no packet loss, and low call dropping probabilities.* In addition, with respect to real-time applications, *minimal buffering* is also included.

For low signalling for micro-mobility movements, *Binding updates are* preferred to be controlled locally, at a cross over-router level, rather than at a home agent.

Chapter Three

LITERATURE REVIEW

3.1 Introduction

This chapter discusses the previous handoff schemes that have been proposed in the literature. The previously proposed fast handoff schemes are discussed relative to how the L3 *handoff is detected and anticipated* as our framework. This expands to: *(i)* history-based mobility profile handoff strategy, *(ii)* location aware and mobility parameter-based handoff strategies, *(iii)* L2 trigger-dependent handoff strategies, and *(iv)* procedure of acquiring or reserving handoff resources. In Section 3.2 handoff management schemes are discussed using the above framework. Section 3.3, discusses the design challenges in L3 handoff. Section 3.4 discusses the generic network layer handoff requirements discovered from the schemes reviewed. Finally, section 3.5 gives a summary of the limitations in the previously proposed schemes that were reviewed.

3.2 Handoff Management Strategies

This research study deals with handoff management in wireless networks. It focuses on L3 (subnet-to-subnet) handoffs. However, since L3 handoff actually includes two handoff layers, namely, link layer handoff and IP layer handoff, a few valuable L2 handoff schemes are also discussed. In the next subsections below, we discuss the previously proposed handoff schemes. We categorise these schemes as follows: *(i)* history-based mobility profile handoff strategy, *(ii)* location aware and mobility parameter-based handoff strategies, *(iii)* L2 trigger-dependent handoff strategies, and *(iv)* handoff resource reservation procedure.

3.2.1 History-based mobility profile handoff Strategy

This strategy uses MN's movement history to predict and initiate handoff. The network or MN keeps records of handoffs occurring, as MN moves across cells. In future, when the MN travels on the same path and needs to execute handoff, the history profile that contains previous handover points is consulted. The main aim of such strategies is to achieve fast handoff by skipping the prediction task. There is also an advantage in that MNs need not wait for beacon messages from surrounding subnet cells before initiating handoff. Typical schemes that use this strategy include (Feng and Reeves 2004, Van den Witjngaert, 2002).

Van den Witjngaert, (2002), asserts that the movement history of a Mobile Node is kept by entities known as Foreign Agents (FAs). Each time the Mobile Node registers with a new Foreign Agent, it tells the new Foreign Agent about its previous Foreign Agent. When the new Foreign Agent sends back binding update to the previous Foreign Agent, the previous Foreign Agent also learns MN's new Care-of-address. This way, each FA can build a table about next hop FAs. Tunnels are also built between FAs. On handover, the FA retrieves MN's history of handovers and starts sending MN's data to the next hop FA in the table, ahead of MN. This scheme gives the advantage that once the channel has been setup; it will be re-used as long as MN keeps coming back in that route. The authors attest that this provides fast handovers. However, this scheme is limited to city streets, where paths are usually straight roads and thus, are so predictive. Therefore, the scheme cannot be effective in a random walk model.

Feng and Reeves (2004), on the other hand, contends that MN records network-layer movements, and stores the IDs of the previously visited subnets. On handoff, the MN uses those records to predict the future target (next subnet). In case the MN has no records (i.e. first time visit) there will be no proactive handoff, so mobile IPv4 will be used. This scheme is said to reduce L3 handoff latency to the level of L2 handoff latency. The binding updates in this scheme are served by a home agent (HA). The limitation of this scheme is that it can only work for MN with previous history. Furthermore, considering the limited capabilities in MNs, the recording, calculations, and prediction can be a heavy burden to the MN. Also there is long signalling delay since HA is used as the signalling anchor point.

In Dealing with call dropping problems, (Kim, 2001) proposed an algorithm that uses user's mobility pattern for effective channel assignment. The user's mobility pattern can be learnt or recorded from the user's daily activities and routes taken on a daily or weekly basis. These are stored in a database. Then the probability that the user can move to a particular cell is calculated, thus channels are reserved accordingly at that particular target cell. This method cannot give precise prediction about changes in the user's speed. The user's speed can change at any moment, thus the recorded details in the movement pattern become partially invalid.

History based handoff strategies are limited by the fact that they are not able to work if there are no previous handover records of MN. Furthermore, historical data itself is problematic, in terms of high storage requirement, and computational costs. They sometimes have misleading information because they do not represent the real-time information about the mobility status of the mobile node. This can also be less effective in terms of supporting speed variation of MNs. For example, if the MN last performed handover at point *x* when it was moving at say 10 m/s, in future when it comes back the same path but now moving at 40 m/s it will be made to still use the same point for handoff. The call carried by that mobile node is now subjected to high chances of it being dropped.

3.2.2 Location Aware, and mobility parameter-based handoff strategies

This approach is usually applied over pre-configured zones in a cell. The zones can be critical regions where the MN highly requires a handover. Choi et al (2000), and (Lu et al, 2004) are typical examples where a cell is marked into zones with respect to handover phases requirements (for example, preparation and execution). The advantage in such schemes is that handoff resource reservation commence well in advance before the handoff execution point.

Choi et al (2000) proposed a Mobility Pattern Profile (MPP) and a 2-tier cell structure for the reservation of bandwidth for hand-off calls in the next cells. “Tier 1” is an area of cell between BS and Tier 2. “Tier 2” is the area close to the cell border, where MN is supposed to execute the handover. This scheme classifies traffic into two Classes: class I (real-time traffic) and Class II (non-real time traffic). Reservations are only made for Class I clients that reside in “Tier-2”. The idea behind the scheme is that the MN’s handoff is recognised with respect to location zone (tier) in a cell. The moment of triggering of hand-off is predicted from Pilot Signal Strength (PSS) received by the MN. Figure 3.2 represent a cell with zones referenced to as “Tiers”.

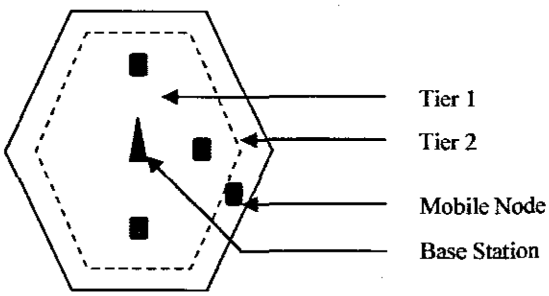

Figure 3. 1 The 2-tier cell structure

Lu et al (2004) proposed a strategy that divides a cell into three zones namely, non-handoff zone –area near the BS; pre-handoff zone – an area starting from non-handoff zone and stretching outwards but before the handoff zone; and handoff zone – an area close to the cell border. When the Mobile Node (MN) is in pre-handoff zone, handoff is anticipated, hence reservation of resources begins. Furthermore, the MN's mobility parameters are constantly checked, every second. When MN is in the handoff zone, handoff can be triggered on request. The advantage of this scheme is that resource acquisition commences in advance, giving enough number of re-tries, in-case resources are currently blocked in the next cell. This technique is currently limited to L2 handovers. If it can be upgraded to accommodate L3 handovers, high speed mobile nodes can get better service in terms of handoff call dropping chances.

Another realistic and popular alternative of achieving fast handovers is to use MN's current mobility profile (movement parameters) in real-time basis. This is aided by the use of GPS (Global Position System). If MN is GPS-enabled, it is easier to determine its position within a cell. The advantage of this strategy is that it gives the real-time mobility profile of the MN, thus has better accuracy in terms of handover prediction, including target cells. This strategy is feasible for both link-layer and network-layer handovers. (Lee and Kang, 2000) and (Xu and Ye, 2002) present L3 handover schemes that are aided by the use of GPS.

Lee and Kang (2000), assume that MN can predict the change of FA ahead of time. The Base Stations are IP routers (FA) capable of handling IP packets. The operation of this scheme is based on an assumption of a free-space where MN's received signal strength is dependent on how far it is to the FA. MN is required to know details of its position and speed through GPS, and then alert the FA about these details.

Xu and Ye (2002), proposed a scheme that uses GPS measurements to determine when channel reservations are to be made. The MN is assumed to be GPS-enabled. Handoff reservations are made based on position, orientation and relative motion of MN to the target cell. MN keeps measuring its coordinates regularly. The BS then tracks MN's previous position and calculates the relative speed with respect to the next cell. Therefore the BS can determine when reservations are to be initiated. The problem faced by such position aware-strategies is that, they are not triggered autonomously. They have to wait until signal strength level indicates a need for handoff. Therefore high-speed-moving nodes cannot be effectively catered for since signal-strength-determined-handoffs always react when MN is close to the cell boundary.

High speeding mobile nodes are problematic when performing handovers. They are affected by late detection of handovers (relative to their speed of travelling), handoff resources acquisition delays, and registration delays. This is worse in L3 handover, since current handoff strategies have no mechanism of anticipating L3 handover without any indications from Link layer.

Fournogerakis et al (2001), proposed a Location-Aided Handover mechanism (LAH). LAH is comprised of a set of algorithms (such as handover prioritisation algorithm). This scheme is driven by the assumption that, it can differentiate mobiles with high-data rate services, and hence the network will be able to predict the next target cell the mobile is likely to handover to, in advance. Hence resources are reserved accordingly. Unfortunately, in this scheme it is not clear as to when reservation starts exactly. Starting reservations too early can block resources and starting too late can cause high handoff call dropping probabilities.

Ivanov and Spring (1995), Sung and Wang (1994) and Iera et al (2002), attempt to solve the problem of very high speed mobile nodes by proposing an architecture that is made up of macro-cells (upper layer) and micro-cells (lower layer). The macro-cell is regarded as an umbrella cell to micro-cells. The slow mobiles from macro-cell are directed to micro cells. High speeding mobiles from micro cell are directed to macro cell. Mobile node's speed is estimated with reference to its dwell time in a cell. If the MN crosses cell boundary a shorter time than threshold, that MN is said to be fast, and vice versa. Their aim is to minimise frequent handovers by assigning high speed mobiles to larger cells.

Hernandez and Helal (2004) gave a different approach to accommodate very high speed travelling nodes. Hernandez and Helal (2004) proposed a handoff scheme with proactive approach that extends MIP by introducing two entities: Ghost Mobile Host and Ghost

Foreign Agent. The ghost-MN is a virtual repeater that always travels with MN along the cells in a determined path. Its functions are to replicate registration requests, Authentication, and Authorization, and create tunnels on behalf of MN. The ghost-FA is created in the neighbourhood FAs and advertises the presence of FA to those neighbour FAs so that MN in those FAs can easily get router advertisements before they request for them. This method can aggravate the shortage-of-resource problem, because those entities will also share the scarce resources. It is also not clear as to how MNs can be served by a single ghost entity.

3.2.3 L2 trigger-dependent handoff Strategies

This strategy makes use of L2 pre-indications about forth-coming link layer handoff. The triggers are closely related to the signal level characteristics i.e. can be triggered when the signal drops below a certain threshold level. Information contained by L2 triggers indicate the start and end of L2 handoff, link state change, i.e. link up, or link down (Kempf et al, 2001). The advantage of using L2 triggers is that the network layer handoff can be initiated earlier.

Fast handovers FMIPv6 (Dommety (2002)), is one of the handover protocols that uses L2 triggers to anticipate L3 handover; hence L3 handoff preparation is started in advance. As the mobile node moves near the cell boundary, the received signal strength value depreciates gradually. When signal threshold value is reached, L2 trigger fires an event that reports impending link layer handoff. The network layer reads information from these triggers and begins to map the prefix of their IP address. If their prefix address does not match its IP layer prefix address, then this implies that the MN is handing-off to a cell belonging to another network. Hence, L3 handoff is anticipated and started before L2 handoff finishes. FMIPv6 was developed with an aim of improving on MIPv6's handoff detection mechanism by starting L3 handoff initiation through an anticipation procedure. It is worth noting that, one of the limitations of Mobile IPv6 Protocol is that, it can only detect handoff after the Mobile Node has moved to a new access router's service area.

FMIPv6 may reduce latency and minimise handoff call dropping rate effectively for slow and medium speed travelling nodes. However, there are some limitations that can deteriorate the service under this protocol. Firstly, currently there is no standard procedure on how triggers are generated. In case the L2 triggers are not available early, then the protocol is delayed. If they are absent, as in the case of non-overlapping networks, the protocol cannot work. In a situation where the MN is moving fast, the tunnel can be extended to the third AR. This therefore implies longer path extension. High speed nodes are exposed to high call dropping probability. FMIPv6 incurs relatively higher signalling delay because binding updates are sent to the HA. Other schemes that use the same principle on anticipation of network layer handoff are presented in (Soliman et al (2002), Jung et al (2005 [a], 2005 [b]), Langar et al (2006), Lin et al (2006), and Saraswady and Shanmugavel (2004)). They also use L2 triggers to initiate their network layer handovers.

(Jung et al, 2005 [a], 2005 [b]) proposed a Fast Hierarchical Mobile IPv6 handover scheme (FHMIPv6). This scheme integrates HMIPv6 and FMIPv6 (Salkintzis, 2004) to achieve fast handovers. This scheme similar to FMIPv6 also uses L2 triggers to anticipate L3 forthcoming handover. The aim of this scheme is to localise binding updates and eliminate tunnelling from previous access router. In this scheme the binding updates are served at the Mobility Anchor Point (MAP), thus this results in low signalling. Instead of establishing a tunnel from previous access router (PAR) to new access router (NAR), a tunnel is established between MAP and NAR. These characteristics offer low signalling and delays. This scheme adds no additional message to the original FMIPv6, except modifying some procedures to facilitate faster handoff. Figure 3.2 shows a diagram for a typical HMIPv6 access network with FHMIPv6 tunnel establishment.

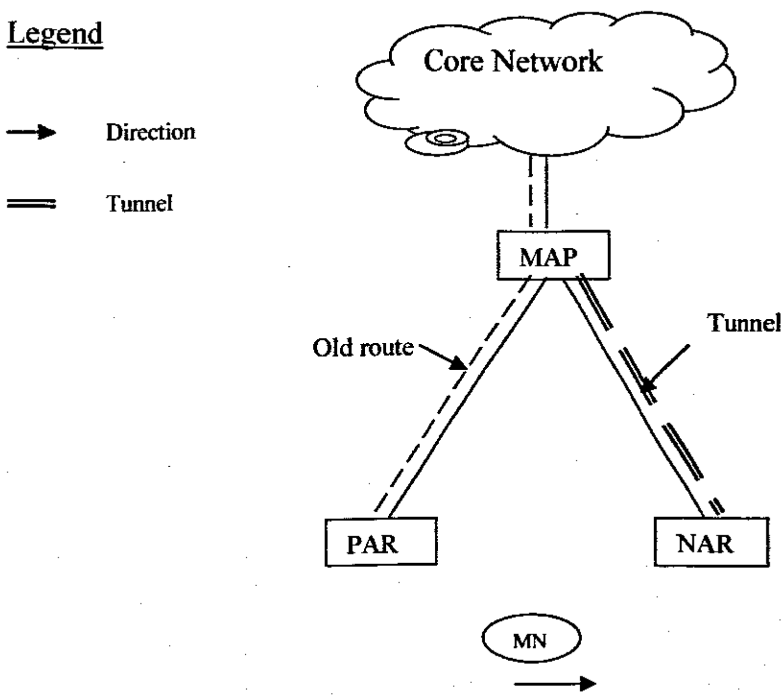

Figure 3. 2 Typical HMIPv6 Access Network with FHMIPv6 Tunnelling

3.2.4 Reservation of handoff Resources

The availability of resources in the handoff target cell provides better chances to the success of handoff calls. However, since wireless networks have scarcity of resources, there must be proper ways of making reservations on time. Handoff resource reservation can be in two ways: *in-advance reservations,* or *immediate reservations.*

(A) In-advance reservations

This technique makes reservations to adjacent cells along MN's path. The advantage here is that when the MN wants to handover to the adjacent cell, the resources are already in place. Thus there is less probability that the handoff call can be dropped due to unavailability of resources in the target cell.

Huang and Chen (2003) and Fournogerakis et al (2001) proposed schemes with advance handoff resource reservations. (Huang, 2003), proposed an RSVP extension with the aim of adapting RSVP to HMIPv6 and provide pre-reservation models for real-time and multi-media services. In their work, an RSVP extension is reviewed under its intra-site mobility architecture aspect. The scheme reserves resources in advance along MN's expected future path. It uses three reservation models: namely: conventional, pre-reservation, and transient reservation. In conventional/RSVP reservation, the MN sends its historical mobility profile to a QoS Agent (QA). The QA then subscribes to the resources accordingly. In pre-reservation, the QA makes reservations to access routers where resources are only reserved but not allocated. In transient reservation, the best effort applications share the pre-reserved resources. When the MN enters a new cell it sends conventional reservations to the QA, and the QA then makes pre-reservation to adjacent access routers. The limitation of this scheme is that its predictive method uses a mobile node's mobility profile that presents the history of movements. Such predictions that are history-based can incur inaccurate results and also do not work for first time visiting nodes.

In some other schemes, reservations are made automatically in all neighbouring cells. The cells form a cluster (group). When MN is given access to one of the cells belonging to this particular group, all the other cells will also receive information meant for the MN. This is advantageous in terms of the elimination of resource requests and registration phases during handover. The main idea in this strategy is that once the MN is authenticated in one cell, the adjacent cells automatically receive MN's context and content, hence reserving resources for the MN.

Tan et al (1999), proposed a strategy where base stations are logically organised into Dynamic Virtual Macro-Cells (DVMs). DVMs are formed by clusters of base stations adjacent to each other. The base stations are network-layer routers with buffers and are capable of subscribing to multicast group. When the MN is served by one BS of the multicast group, all the Base Stations in the multicast group receive MN's information. This makes it easy, such that if the MN moves to an adjacent BS, the MN's recent information will be forwarded immediately. But there are some limitations in the scheme, such as heavy buffering, where all adjacent cells have to buffer packets. This means that they could create unnecessary congestion and blocking of resources. There is no predictive measure to predict future handoff. It only relies on beacon messages from neighbouring subnets, which usually take several seconds interval. This way, due handoffs can be delayed, and there can be high call dropping probability for fast travelling nodes. (Helmy, 2002), and (Lo, 2004), also made use of a similar concept.

Pack et al (2005), proposed a Selective Neighbour Caching Scheme. It calculates neighbour graph and neighbour weights probability from the statistics of the handoff patterns amongst the APs, recorded in a weight table. The MN's context is then proactively propagated only to the selected neighbours with higher weights. This scheme has a better selective multicasting strategy where only cells with high probabilities can receive the MN's context, thus less resources occupied. However, it is not specified as to when the context is propagated.

In the quest to provision fast handoffs, some other schemes further transfer the security context in advance so as to eliminate security association delays. Samprakou et al (2005), Chou and Shin (2005), and Duong et al (2005) proposed a concept of transferring MN's security context in advance to surrounding FAs. This minimises handoff duration because authentication and authorisation have been carried out earlier.

(B) Immediate reservations

In this strategy, handoff resource reservation starts when handoff commences. The handoff schemes that fall under this category, are usually signal-measurement based. They are realised when the signal strength from current BS falls below the one in the next BS, thus initiating handoff.

This method has the potential for high handoff dropping probability. With soft handoff, it should be noted as well that, when handoff is initiated, it does not jump to the switching point (changing point of attachment) stage. There are conditions that need to be satisfied first; for example, have resources been granted by the target cell? If not, there is no way the switching process can continue. Let us take into consideration a situation where the mobile node is moving very fast and there is only a little overlap region between cells. It may happen that on handoff initiation, the resources are not available yet, in the next cell; so the request has to wait. While the request is still in the queue, the mobile node (due to its high speed) may move deep into the new cell, completely losing signal from the old cell. The call then will be forced to be dropped.

3.3 Network Layer (L3) Handoff Design Challenges

Network layer handoff is a more complex type of handoff than L2 handoff because it involves two layers, namely: link-layer, and IP layer. Therefore, there should be a proper mechanism of detecting and synchronising the occurrences of handover among these layers. Although the common goal of the schemes reviewed in section 3.2 is to provide fast and robust handovers (minimising latency, and call dropping probabilities), they are still deficient with respect to the following characteristics: 1) Early L3 Handoff Recognition mechanism, 2) Mobile Node's Speed-aware handoff preparation and execution decisions, 3) Enhanced prioritisation in handoff queue (strategic access to handoff resources). These are elaborated upon in the subsections that follow.

3.3.1 Early L3 Handoff Recognition mechanism

Early anticipation of L3 handover is aided by some indications that can be gathered from the lower layer (Link-layer). These indications could be in the form of L2 triggers which when detected can be mapped to identify their network prefix address (Trossen et al ,2001). These triggers are not autonomous. They are triggered when link layer shows some indication of fading. The free-space prediction mechanisms too, that may use MN's real-time mobility parameters also require some indications about current and future status of Link-layer service (signal levels) in order to start anticipation of L3 handover. These indications usually come with link-deterioration state which comes when the MN approaches the cell boundary. Considering reduced cell sizes in wireless networks and high speed travelling MNs, this means that calls carried by such MNs are subjected to high probability of being dropped.

3.3.2 Mobile node's speed-aware L3 handoff decisions

When L3 handoff has been anticipated through triggers or any other means, preparation immediately begins. The objective is to finish L3 handoff closely to the finishing of L2 handoff (Samprakou et al, 2005). The MN's speed is not further closely monitored. This could result to some worthless proactiveness. For example, if the MN suddenly reduces speed, there is a possibility that L3 handoff finishes even before L2 handoff starts, or possibly proceed with L3 while MN has suddenly become stationary.

3.3.3 Enhanced prioritisation of handoff queue

Generally the priority of accessing resources at the next cell, for a handoff call, is usually based on the class type of the application running on the mobile node, its position to the boundary and received RSSI, (Abdulova and Aybay, 2006) and (Choi et al, 2000). Real-time applications are treated with high priority. Consider a case where two mobiles, M_1 and M_2 carrying calls of the same class (say real-time applications), are at the state of executing a handoff. M_1 is in front of M_2 by x distance, moving at lower speed than M_2, and its resources have already been reserved. The system is currently at its full state. M_2 starts requesting, and by calculations M_2 will pass M_1 at some point soon. Since the two mobiles are in the same zone, using first-come-first served algorithm, because M_1 reserved first, M_1 will continue holding its share, thus M_2's call has to be dropped. If the relative motion (speed, position, and location; more especially the speed) of each MN to the boundary of the current serving cell was to be considered, such a scenario could be resolved by letting slower mobiles give up their share to faster mobiles, thus leading to handoff queue being reshuffled. Unfortunately speed comparison is hardly considered.

3.4 Generic handoff requirements

In dealing with network-layer handoffs, most of the proposed strategies in the literature have adopted hierarchical architectures as a means of localising micro-mobility handoffs and thus reducing signalling overheads for route updates. In supporting fast handovers, the proactive approach is the most favourable applied technique. This minimises handoff latency, and call dropping probabilities (for average speeding mobile nodes). In Fast handoffs the proactive approach is achieved by the use of L2 triggers to anticipate L3 handoff, hence starting the handoff preparation in advance. The other means of providing fast handoff is by the use of MN's history of handover points as a means of detecting due handoff and guessing the next cell the mobile node might move to. The strategy of sending packets in advance to the predicted future locations is highly favoured and is a commonly adopted procedure in order to reduce packet loss during handoff.

3.5 Summary of limitations of previously proposed schemes

The previously reviewed schemes have the following limitations: *(i)* Use of historical data as means of anticipating and predicting handoff. Historical data is not usually accurate enough. It cannot work for mobiles without any previous records about their past movements in that particular cell. *(ii)* Use of L2 triggers is impractical where cells do not overlap. They are not autonomous, since they come with signal deterioration. Thus high speeding mobiles may not be well supported (due to smaller cell sizes). *(iii)* Their buffering rate tends to be too high for real-time applications.

The limitations of the previously proposed layer 3 handoff strategies make them less efficient in terms of supporting real-time applications and high speed mobile nodes. As an attempt to providing support for real-time and non-real-time applications, but more especially *the real-time application and variable speed mobile nodes*, we propose an L3 (IP layer – inter-subnet) handoff strategy that uses MN's movement parameters (location, speed, and direction) in real-time over confined cell areas (cell sectors). We make use of a mobile node's movement parameters in real-time as a means of anticipating L3 handover; thus triggering a prediction mechanism in order to determine the target access router. We explicitly separate handoff into two phases, viz: preparation and execution phases. The main idea is to modularise the handoff process into units occurring at appropriate phases. Detailed description of the proposed strategy follows in the next chapter.

Chapter Four

MODEL DEVELOPMENT

4.1 Introduction

In the previous chapter we presented some already proposed network layer handoff schemes. However, these schemes have several limitations, as far as real-time applications' QoS requirements and speed variation of MNs are concerned.

In order to address some of these limitations and challenges, in this chapter, a Sector Aware Handoff (SAH) scheme is proposed. The SAH offers a better approach in anticipating and recognising L3 handovers. In this scheme, speed variation can be supported effectively, thus minimising call dropping probability. The chapter also describes the architecture used by SAH and the handoff procedure employed in the scheme.

4.2 SAH Design Goals and Principles

Handoff schemes that anticipate L3 handover through the use of L2 triggers are relatively fast and reduce handoff latency.

SAH's design goals are similar to those mentioned in Chapter 3, which are to provide: a) Early anticipation of L3 handover; b) Mobile Node's speed-aware handoff preparation and execution decisions; and c) Enhanced prioritisation in handoff queue (i.e. strategic access to handoff resources). These are elaborated below.

(a) Early anticipation of L3 handover

In the SAH scheme, a fast handoff is achieved by making use of cell sectors. Prediction is also incorporated in order to anticipate L3 handover early enough. In prediction, the next-cell prediction scheme as proposed by (Nkambule et al, 2004) is employed. This scheme partitions a cell into sectors (six sectored cell), and within each sector are two regions: (i) Non-critical region – area near base station; and (ii) Critical region – area after non-critical region up to the boundary of the cell. The prediction is triggered when the MN enters the critical region. The scheme uses a mobile node's movement parameters, i.e., speed, position, and direction that present current mobility information.

(b) User's speed-aware L3 handoff decisions

SAH makes use of demarcated zones for each phase in handoff. In addition, it also takes into account the real-time movement parameters of MN in its decision making process. In the SAH scheme, the preparation zone is made larger for fast moving mobile nodes, and smaller for slow moving mobile nodes. In this way, fast moving mobile nodes will be served earlier than slow moving nodes situated around the same position.

(c) Enhanced prioritisation in handoff queue

In SAH, in addition to MN's position to the cell boundary or target cell, the real-time speed of MNs is also taken into account, thus slow mobiles would be given lower priority than fast mobiles. For an example: consider two mobiles m1 and m_2. Assume m_1 is moving slowly and has requested for handoff and has been placed in the handoff queue (as the case when resources are not available yet). Also assume M_2 is moving faster than m_1 and are in the same position. In the handoff request queue, m_2 is behind m_1. Due to speed consideration, the handoff queue requests would be reshuffled accordingly provided that m_1 by calculation would not be affected (i.e. increase the probability of getting dropped).

4.3 The Proposed Model

We adopt a hierarchical architecture (Soliman et al, 2002) and provide enhanced L3 handoff anticipation mechanism. The components in our model are: Domain Agent (DA), Subnet Agents (SAs), Access Points (Base Stations), and virtual demarcated Cell Sectors. Figure 4.1 represents the architectural model and cell clustering scheme. This structure has also been used in Sung et al (1994).

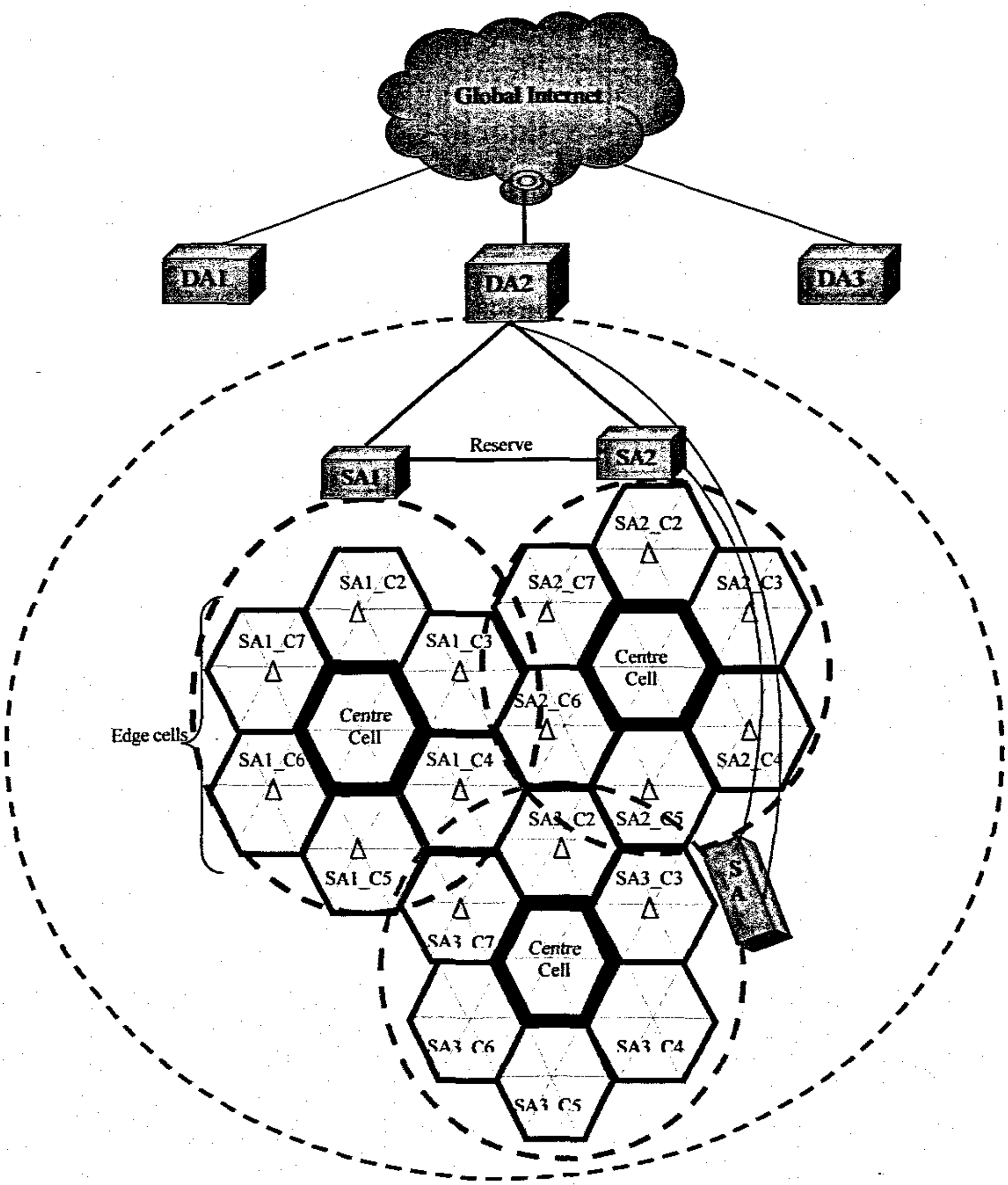

Figure 4.1 Inter-subnet handoff design architectural model

4.3.1 Domain Agent (DA)

The DA is a router controlling several subnets, and acting as a gateway for a group of subnets. In figure 4.1, the DAs are DA1, DA2, and DA3. The DA, in our model, performs the following functions:

- Hides MN's mobility within its domain and administers location updates as long as the MN is under its service area. Act as binding updates control point;
- Routes traffic to respective destination subnets. It also carries out requests from SAs with regard to path switching during handoff;
- It keeps MN's entries (routing table) regarding subnet location and updates those entries as per network layer handoff and
- During handoff, duplicates and bi-casts packets to both the old SA location (where MN is preparing to leave) and the new identified target SA.

4.3.2 Subnet Agent (SA)

A subnet Agent, in this architecture, is a router that manages a cluster of seven cells. It is similar to access router (AR) in functionality. A larger grouping e.g. 9, 12, 19, and 21 can also be utilised but 7 is considered the minimum as in our model. In this grouping, the six cells surrounding the centre cell are referred to as *edge cells*. During handoff, the base station serving the MN, reports to its SA. The SA studies the nature of handoff being requested. In case the handoff involves IP layer, SA passes the request to DA.

The L3 handoff in our architecture occurs when the mobile node moves from a cell belonging to one subnet to a cell belonging to another subnet e.g., from subnet-1 to subnet-

2 (SA1_C3 to SA2_C6) in Figure 4.1. This is called inter-subnet mobility. There is also an intra-subnet mobility, which occurs when a mobile moves between cells that belong to the same subnet, e.g., from SA1_C7 to SA1_C2. Therefore, our model has two levels of mobility i.e., intra-subnet and inter-subnet mobility. The main focus of this research is on inter-subnet mobility handoff.

4.3.3 Demarcation of Subnet cells

The proposed model takes into consideration the difference in the span area from two technologies: (i) *Wireless IP-Based Cellular networks* and (ii) *Wireless Local Area Network (WLAN)*. In IP-Based Cellular network, a base station's transmission range has a larger span area of several kilometres (typical radius of 1.6 km – 19 km) (Forouzan, 2003). The WLAN Access Point on the other hand has a shorter range of transmission (below 100 m) (Kim, 2001). The proposed solution makes use of the Wireless IP-Based Cellular Networks architecture. The details with regard to this architecture are described in the next sub-section.

In IP-Based architecture, we demarcate regions in the cell sectors. This scenario is achieved by varying the size of the handoff preparation and handoff execution regions. The edge cell sectors are marked into three regions: (i) Non-handoff region, (ii) Handoff Preparation Region (HPR), and (iii) Handoff Execution Region (HER). The non-handoff region can be consumed by HPR region as HPR increases its size. The most important regions are HPR and HER. These regions are meant for handoff preparation and execution phases respectively. The basic idea is about modularising handoff. It is also to be noted that the preparation phase can last until the MN enters the execution region, in case it cannot finish inside the preparation region. This entirely depends on the MN's speed, and resource availability. If resources are not available the request keeps persisting till a call dropping decision is announced. It is worth noting that, one of the design goals is to accommodate MN's speed variation. Figure 4.2 gives an illustration of regions in the cell sector approach. This idea of concentric regions has also been implemented in Choi et al (2000), Lu et al (2004) and Chung et al (2005).

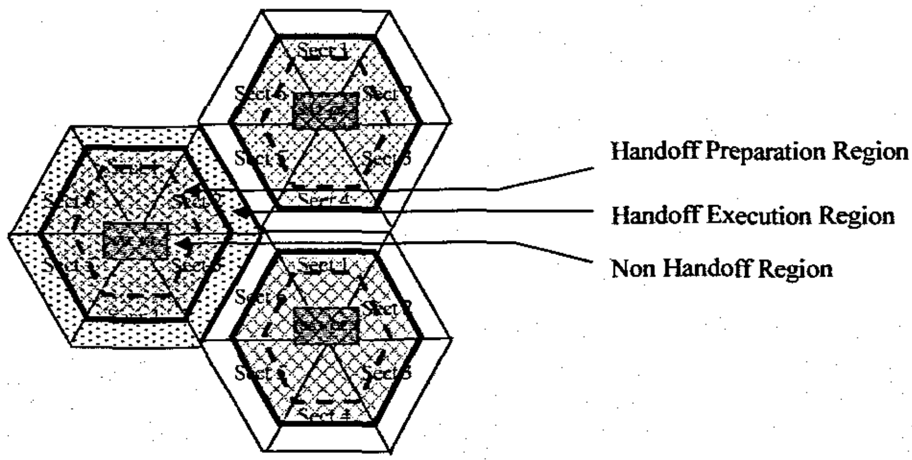

Figure 4.2 Regions in Sector

Figure 4.2 shows three edge cells from different subnets i.e., SA1_C4, SA2_C6, and SA3_C2. Each sector has three regions as shown in SA1_C4. The size of these regions can be varied (made smaller or larger).

4.4 Handoff Procedure

Our handoff occurs in a proactive manner. The subnet-to-subnet level of handoff can be regarded as a combination of two levels i.e., link layer (L2) (change of radio link or access point) and IP Layer (change of Care-of-Address). Our approach (similar to the Fast Handover algorithm (Dommety et al (2002)) acquires information (e.g., IP address) about the next subnet, in advance. The details of the next subnet are gathered through the neighbour casting protocol (Shim, 2001), (Park, 2006); hence prediction can be triggered at any point within the cell when handoff is required.

For intra-subnet movement, a scheme such as that proposed by (Lu, 2004) could be adopted. In this scheme, the cell area is marked into zones, with respect to handoff phases. There are three zones: (i) non-handoff zone, (ii) pre-handoff zone, and (ii) handoff zone.

For inter-subnet movements, we apply our proposed scheme. In this work, handoff is explicitly divided into two phases: *preparation phase and execution phase*. There is a "ready state" after the completion of the first phase and MN is notified and saves the state. Our proposed solution also does not deviate from the basic principle that, signal characteristics is a major component in deciding the handoff switching point. Although the execution phase is in a designated region, signal characteristics are still incorporated in the execution decision process. The detailed description of these two phases is provided in the following two subsections.

4.4.1 Handoff Preparation Phase (HPP)

Handoff preparation occurs in the HPR region. It may last until the MN enters the execution region, depending on two very important parameters; MN's travelling speed, and status of resource availability in the target cell. The main activities during this phase include, but are not limited to: (i) L3 handoff Prediction, (ii) next subnet-router prediction, and (iii) request for resources, including new CoA establishment.

A) L3 Handoff Prediction

Network layer handoff prediction is one of the most important components in providing fast L3 handovers and low call dropping probability. Early anticipation of L3 handover enables early identification of handoff target cells where the node might move to. Early identification of handoff target cell(s) provides the opportunity for more attempts in terms of resources requests (reservations). This is applicable in cases where all available channels in the target cell(s) are currently occupied during the request time.

To provide early L3 handoff prediction, SAH adopts the concept proposed by Shim et al (2002), and Park et al (2006). SAH combines this concept with free-space prediction aided by GPS over a cell sector. These schemes allow a subnet cell to know the details about its neighbour subnets. In building up a table of adjacent subnets, signals are read from mobiles that are near the sector boundary. These mobiles receive signals from other cells, belonging to other subnets. Thus the adjacent subnets' IP addresses can be extracted from these signals. Then each SA can build up a table containing information about adjacent SAs. The sector plays an important role here by allowing geographical location identification of those adjacent subnet cells. Each subnet-cell knows about the geographical location of its sectors. For an example: sectors can be identified using cardinal points, with the BS as the reference point e.g., sector-1: North, sector-2: North East, etc. On L3 handoff prediction, the SA uses this information to anticipate the next subnet router. The availability of such data in the table enables the L3 handoff prediction to occur at any point within the cell as long as the MN's direction is known. The mobile does not necessarily need to get some signals from the next cell. The only information required is: mobile node's *location, direction, and speed.* Table 4.1 shows a typical format of the information table.

Table 4. 1 SAs typical information table

Subnet id	*Cell id*	*Sector id*	*Sector id*	*Cell id*	*Subnet id*
SA1 Network Address	SA1_C3 Mac: Address	Sector2 North East	Sector5 South West	SA2_C7 Mac: Address	SA2 Network Address

Current SA Adjacent SA

Using the information table as given in table 4.1, gathering of adjacent SA's information is simpler. The naming conversions (SA, SA2_C7, Sector2, etc) are referenced from the model.

B) Target cell Prediction

The next subnet cell prediction follows immediately after the anticipation process. (Nkambule et al, 2004) is employed as our prediction scheme. The main task here is to identify one, or two possible subnet cells that have high probability that the MN might move into, depending on its direction. Hence, reservations start early if deemed so. Since the information, including location details, of adjacent subnet cells is readily available, the prediction process is simpler. It is assisted by the data in Table 4.1. When the target subnet cell has been predicted, then the resource request begins.

Summary of handoff preparation:

When the MN is in the HPR region, the serving BS checks MN's movement parameters (*position, velocity, and direction*) in real-time, probably by means of GPS. If the velocity is high and direction is towards HER regions, BS sends a notification message to the serving SA. The serving SA, then uses the information on its table about adjacent subnets to predict the next SA from the HER region, where the MN is heading towards, according to the given direction.

When a target subnet is found, the serving SA sends a resource request message to the predicted SA(s). The messages specify all the requirements, and MN's ID. The predicted SA (s), then send a reply message indicating whether the request can be met or not (accept, or reject the request). If the resources are granted, then the serving SA (on behalf of MN) initiates a L3 pre-registration process. The process then finishes by the predicted SAs creating new Care-of-Addresses (CoA). The serving SA then forwards a notification message to both the DA and MN updating them about the forthcoming L3 handover event. This message contains the predicted SAs IDs and the new CoAs for the MN. Both DA and MN keep this information for later reference, during the execution phase. At this stage, the handoff preparation phase is considered completed and the system is ready for the execution phase. Figure 4.3 shows the sequence diagram of the handoff preparation phase.

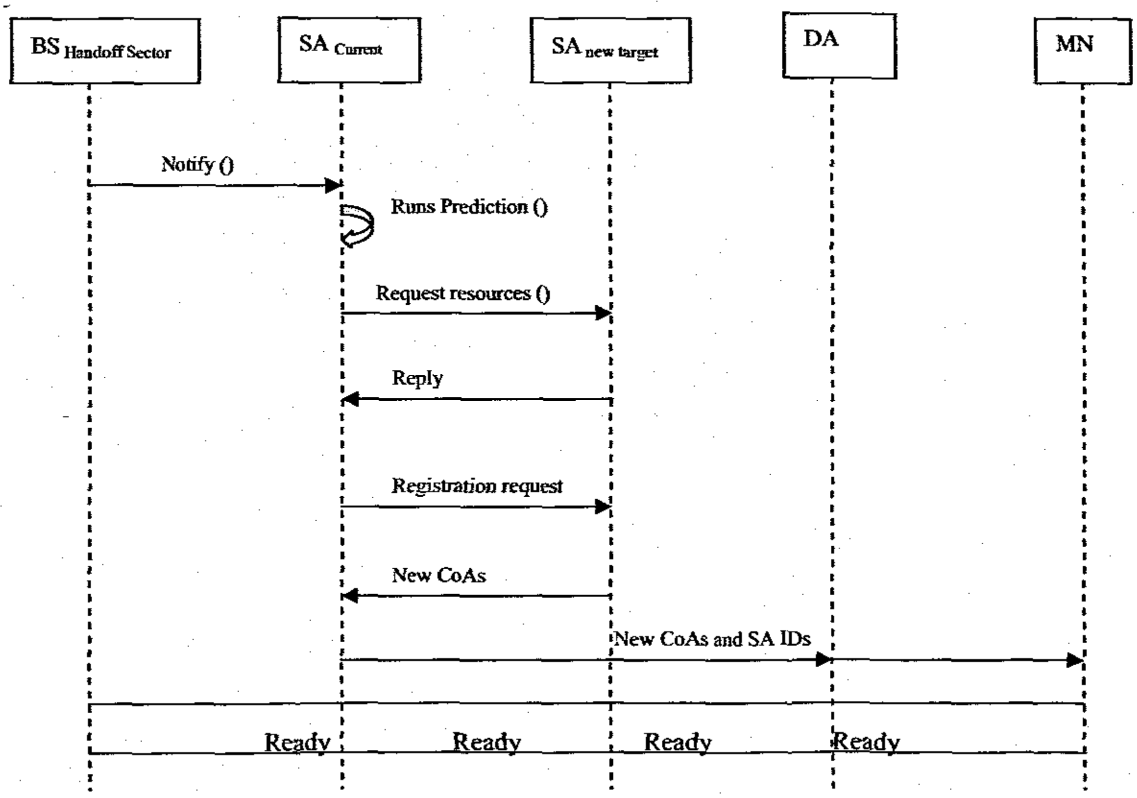

Figure 4.3 Sequence diagram of Handoff Preparation Phase

4.4.2 Handoff Execution Phase (HEP)

When the MN is in the HER region; the serving BS keeps monitoring its behaviour until handoff triggering decision is made. When L2 handoff triggers, L3 handoff is delayed until L2 is almost finish, then the serving BS immediately sends a notification message to the serving SA. This message acts as an update and also indicates the primary target cell. The serving SA, issues a packet duplication (bi-cast) request to the DA. The DA duplicates packets to both the serving SA (old SA) and the predicted SAs (the new target SA and sub-target SAs). The predicted SAs buffer a few most recent packets. As soon as L2 handover finishes, the new BS notifies its SA (the new target SA) to forward packets meant for MN beginning with those in the buffer. The MN sends binding updates request through the new

SA. The new SA passes binding updates to the DA. DA manages the binding updates and redirects the path to point to the new SA. It is to be noted that the HEP phase occurs in parallel with L2 handover. Delaying L3 handoff for some time, helps reduce buffering in the target SA, and this is safe since most of the work has been covered during the preparation phase. A sequence diagram for the handoff execution phase is as shown in Figure 4.4

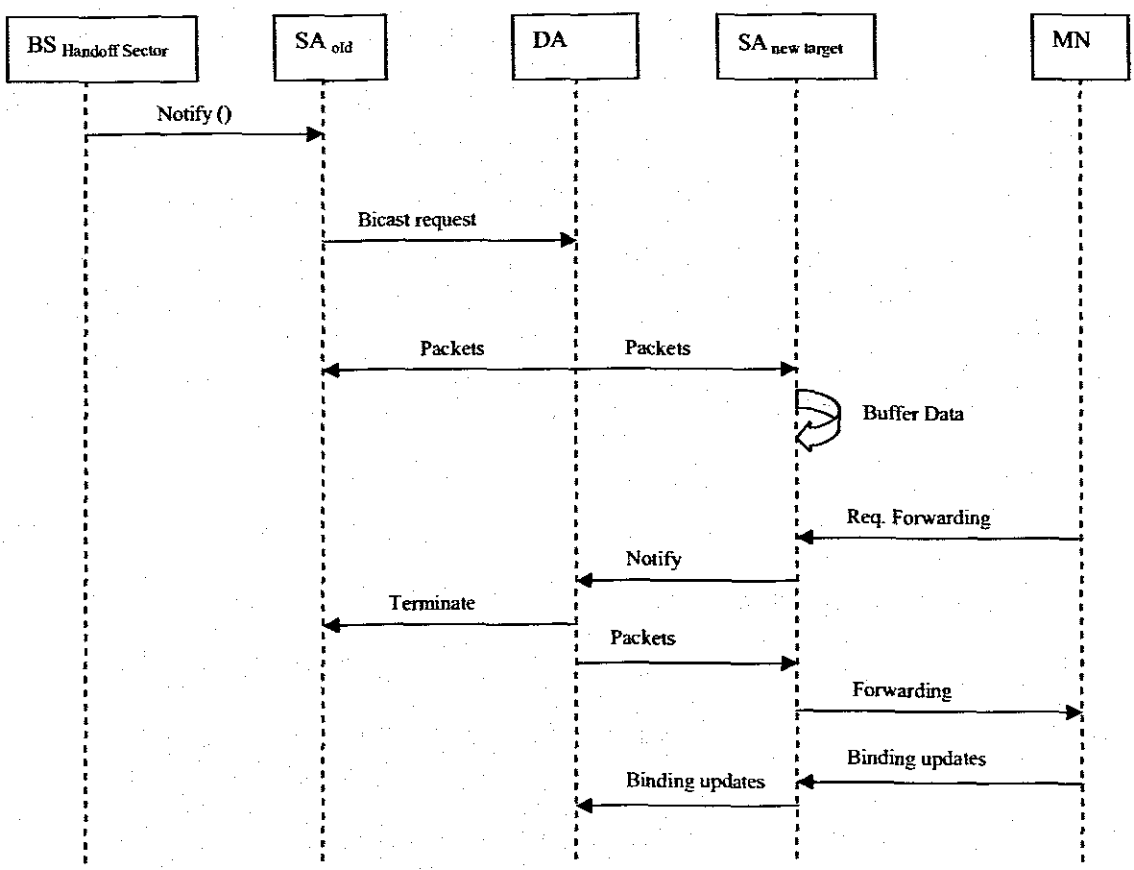

Figure 4.4 Sequence diagram for Handoff Execution Phase

At the completion of this phase, the L3 handover is considered finished. The next subsection provides pseudo code algorithms for the SAH. It details the activities in the handoff procedure using pseudo codes.

4.5 SAH Algorithms

The next two sections present the algorithms used by SAH when predicting L3 handoff. They show how the preparation and execution phases are carried out. The sections explain the procedure for a handoff call. In the final section, the algorithms for admission of calls (handoff calls and new calls) are presented.

4.5.1 SAH Preparation Phase Algorithm

Figure 4.5 and Figure 4.6 present the algorithm for the preparation phase.

HPR	: handoff preparation region sector
HER	: handoff execution region sector
MN (*i,j,k*)	: MN's mobility parameters
	i is direction
	j is position
	k is speed
Predict (*i,j,k*)	: predict using mobility parameters
Prepare (*m*)	: preparation phase for current mobile node
Execute (*m*)	: execution phase for mobile node
Delay_prep(*m*)	: delay preparation phase

Figure 4.5 Pseudo-code notations for preparation and execution of handoff calls

When the system notices a non-stationary MN over the HPR region, the Base Station collects its active mobility profile and passes it to the Subnet Agent (SA). The SA then checks on the speed parameter and decides whether to urgently attend to the alert or to suppress it. If it decides to attend to it, it first predicts the neighbouring subnet that the MN might move to. When the target SA has been identified from the prediction, the preparation phase is initiated (resource acquisition and registration is done). If the resources are not granted, it will keep retrying the requests until it comes to a point where the MN is inside the HER region where it is supposed to execute handoff. If the mobile node's speed is low, the preparation phase is delayed and a timer *t* is started. When the timer *t* expires, the delay is terminated, and preparation starts.

4.5.2 SAH Execution Phase Algorithm

Figure 4.7 presents the algorithm for the execution phase.

```
Pre-Condition: MN moving in HER region
        If MN_handoff_triggered = True then
    Check (MN_Handoff_Status);
            If (handoff_status = ready) then
                Execute (m)
            Else
                Drop (m)
        End if
    End if
Post-Condition: handoff call successful or dropped
```

Figure 4.7 Pseudo-code of execution phase algorithm

When the MN is inside the HER region, the system reads the status of the preparation phase from MN. If the status is "ready", then handoff is executed. If it is "not ready", the handoff call is dropped.

4.5.3 Call Admission Algorithm

This section presents algorithms used by SAH when admitting handoff calls and New calls. In prioritising handoff calls, while also allowing new call to gain access, for simplicity, we employ the guard channel (GC) scheme algorithm (Kim and Jung, 2001), (Abdulova and Aybay, 2006). The total cell's bandwidth is divided into two parts; a part for handoff calls only (Guard Channel) and the other part is shared by both handoff calls and new calls. In the shared part handoff calls and new calls compete for available bandwidth. Handoff calls that could not gain access to the shared part, are redirected to the handoff part. In terms of size, the shared part is more than the handoff part.

SAH employs a simple call admission control criteria that prioritises handoff calls over new calls. When a call comes, the system checks whether there are enough resources in the shared part to satisfy the request. If bandwidth is enough, then the call is *accepted*, otherwise the system further checks whether it is a new call or handoff call. If it is new call, it *rejects* the call; otherwise, if it is handoff call, it redirects the call to the handoff call only, part.

In the handoff part, the system checks whether it is a real-time or non-real time call. Before accepting non-real time calls the system first checks whether there is no real-time call waiting. If there is any waiting then it is served first. A non-real time call can be accepted even if the system cannot meet its bandwidth requirement, as long as the system has some available bandwidth. Figure 4.8 shows the notation used for the CAC algorithm. Figure 4.9 and Figure 4.10 illustrate the call admission control algorithms for the shared part and handoff calls only part respectively.

$Call_{reqst}$	: call request
CallType	: type of requesting call
B_{tot}	: total bandwidth of a Cell
B_{req}	: requested bandwidth
B_{rem}	: remaining bandwidth
B_{used}	: currently used bandwidth

Figure 4.8 Pseudo-code notations for calls

Pre-condition: call requesting

Process Body: Check (*bandwidth*);

$B_{rem} = B_{tot} - B_{used}$;

If $B_{rem} > B_{req}$ then

Accept Call ()

ElseIf $B_{rem} < B_{req}$ then

Check (*CallType*);

If Call = New Call

Reject Call ()

Else

Redirect Call ()

End if

End if

End Body

Post-condition: new call or handoff call accepted OR new call rejected and handoff call redirected.

Figure 4.9 Pseudo-code for the admission of calls in the shared part

```
Pre-condition: Handoff call requesting

Process Body:   Check (Class_Type)

                    If ClassType = Non-Real-time Then

                      Check (if there is any Real-time call)

                        If Real-time = True

                          Process (Real-time Call)

                        Else

                          Process (Non-Real time Call)

                        End if

                      Else

                    Process (Real-time Call)

                End if

End Body

Post-condition: Real-time handoff call processed before Non-
real-time handoff call
```

Figure 4.10 Pseudo code for handoff call admission in handoff part

Chapter Five

SIMULATION, RESULTS AND ANALYSIS

5.1 Introduction

In the previous chapter, the SAH model was proposed. This chapter discusses the implementation of the proposed model. It also provides simulation results and their interpretation. Section 5.2 discuses the implementation, simulation model, simulated environment, performance metrics, and simulation parameters used. Section 5.3 presents the experimental test cases with results presented in both tabular and graphical form. The results are also analysed in this section. Finally, Section 5.4 compares SAH and FHMIPv6 using our performance metrics, followed by discussions on the comparative results.

5.2 SAH Implementation

5.2.1 Simulation Environment

The simulation environment adopted is similar to the HMIPv6 architecture (Soliman et al, 2002). The network entities are hierarchically organised. The domain access network has the following entities: One Domain that consists of three subnets, viz: subnet 1, subnet 2, and subnet 3. The domain is controlled by a Domain Agent (DA). Subnets are controlled by Subnet Agents (SAs). The DA links SAs, and also links the domain access network to the core network. The network uses seven cells per subnet. For clarity, a close-up of a small section has been given in Figure 5.1., where only one edge cell per subnet is shown. Each cell has six sectors and one Base Station. The naming convention is such that each cell name indicates the subnet and the cell in which a Base Station belongs to. For example, SA1_Cell3_BS indicates that it is the base station of cell-3 in subnet-1. Figure 5.1 is an illustration of the simulated network environment.

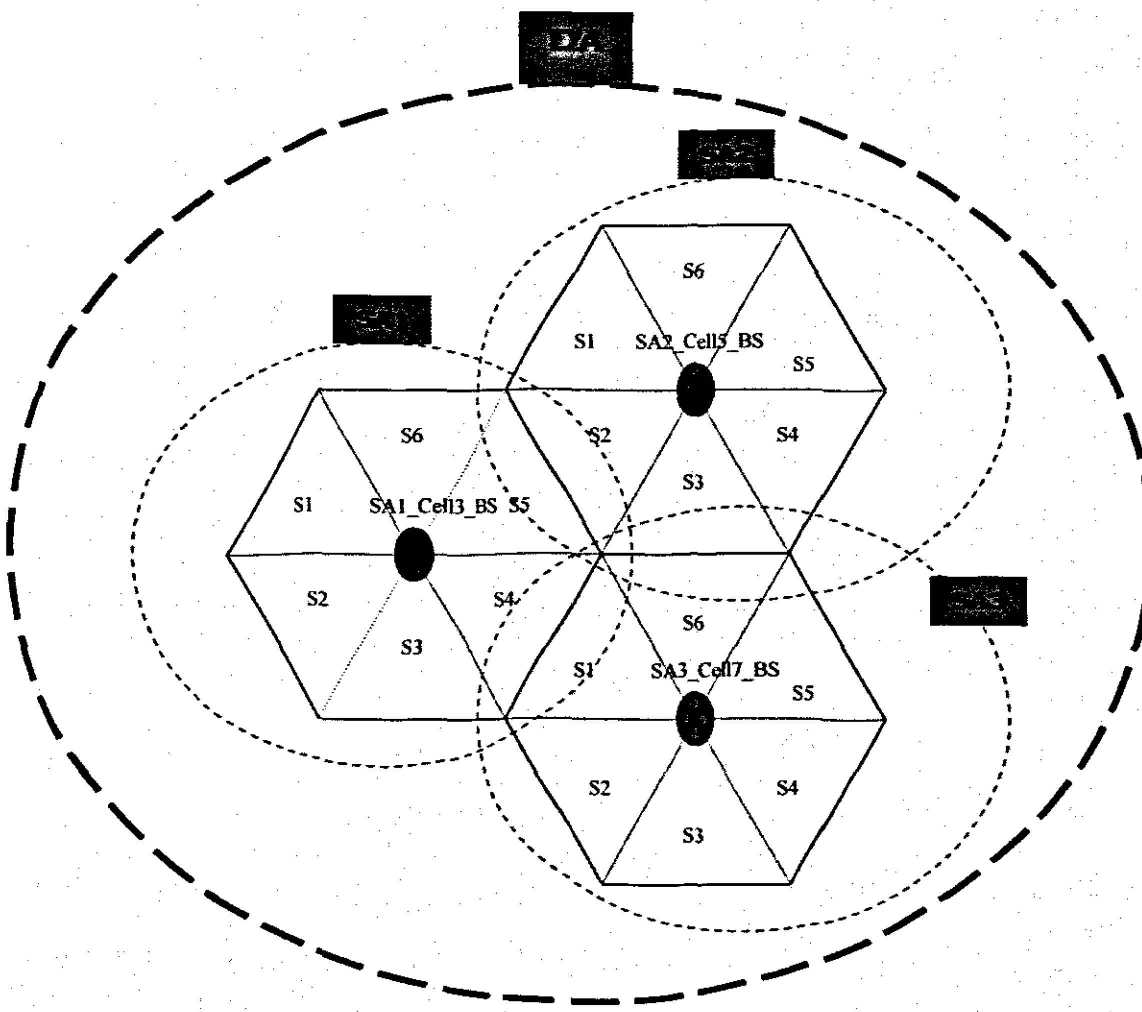

Figure 5.1 Simulated Network Environment

The simulation was developed by using Microsoft Visual Studio.Net 2003 package. The language used from the package was Visual Basic.net language. The Visual Basic.Net language was chosen for being a fully object oriented programming language and easily deployable. Microsoft Access 2003 database was used to store data generated through simulation. The simulation ran on an Intel Pentium processor PC with 3.2 GHz speed, with Windows XP as the operating system.

The objects in the simulation were the Network, Domain Agent, Subnet Agents, Mobile Nodes, and Cells (with sectors, and Base Stations). These exchanged control messages, sending and receiving IP packets. The detailed class diagram is found in the Appendix.

5.2.2 Simulation Model

There were seven cells in each of the three subnets. In Figure 5.1, only one edge cell from each of the three subnets is depicted, for clarity. The cells shown are cell-3 of subnet-1 with base station SA1_Cell3_BS, cell-5 of subnet-2 with base station SA2_Cell5_BS, and cell-7 of subnet-3 with base station SA3_Cell7_BS. The six sectors in each cell are labelled S1 to S6. The maximum number of mobiles was 300. Their movement pattern was random. This allowed them to move around the subnet cells in any direction. The total capacity of each cell was 150 Kbps. Handoff calls were allocated a Guard Channel (GC) of 50 Kbps per subnet cell. The other amount of 100 Kbps was shared by both new calls and handoff calls. When the MN generates a call, it is named a new call. When handoff is initiated for that particular call, it becomes a handoff call. A mobile node was made to generate a call at anytime. Calls could be generated in any region of a subnet cell. Calls were generated at random intervals; from 2 - 20 calls/sec. Their life-time was also random between 15 and 180 seconds. The speed of mobiles varied from low to high. The low speed was taken as 5 m/s while high speed was 40 m/s. These speeds correspond to practical speeds of 18 km/h and 144 km/h respectively. The handoff region also varied from small to large i.e., from 20 % to 80 %.

5.2.3 Performance Metrics

The performance metrics used in the simulation were (i) Handoff call dropping probability (HDP) – the ratio of the number of handoff calls dropped to the total number of handoff requests, (ii) New call blocking probability (CBP) – the ratio of the number of new calls blocked to the total number of new calls attempted, and (iii) Utilisation – the amount of bandwidth used at time *t* compared to the total amount of bandwidth allocated to the cell, expressed in percentage.

The above three performance metrics are presented in an equation form as below.

$$HDP = \frac{H_{dc}}{H_{th}} \quad \text{.......... Equation 5. 1}$$

Where: HDP = handoff dropping probability

H_{dc} = number of handoff dropped calls

H_{th} = total number of handoff calls

$$\text{CBP} = \frac{C_{bc}}{C_{tn}} \quad \text{.......... Equation 5. 2}$$

Where: CBP = call blocking probability

C_{bc} = number of new blocked calls

C_{tn} = total number of new calls

$$\% \text{ Utilisation} = \frac{B_{used}}{B_{tot}} \times 100 \qquad \text{Equation 5. 3}$$

Where: B_{used} = used bandwidth

B_{tot} = total bandwidth in the cell

5.2.4 Simulation Parameters

In evaluating the performance of our system, we used three parameters i.e. *Speed, load, and the size of the handoff region (i.e., the preparation region).*

a) Speed

The mobile node's speed was varied, starting from low speed, 5 m/s to high speed, 40 m/s, which corresponds to 18 km/h to 144 km/h. The behaviour of the system was analysed with respect to the performance characteristics.

b) Load

$$\rho = \frac{N_c}{t} \qquad \text{Equation 5. 4}$$

Where: ρ = load

N_c = number of calls arrived

t = instant time

Load comprised the total number of handoff calls and new calls per second per subnet cell. The rate of calls requested varied increasingly from three calls per second to fifteen calls per second. Again, the behaviour of the system was also analysed using performance metrics. Equation 5.4 shows the load formula.

c) *Size of Handoff Preparation Region (HPR)*

The size of the HPR region was varied from small size to large size (10 % to 80 %). The behaviour of the system was also analysed in order to observe how this change affects the system. The HPR region is calculated as a relative value, in percentage, compared to the size of the cell. This value is obtained by taking the ratio of the length of the small section of radius in the HPR region to the radius of the cell. The HPR region was increased inwards of the cell, starting at the cell boundary. Figure 5.2 depicts a cell sector with an indication of the distance measure of the HPR region.

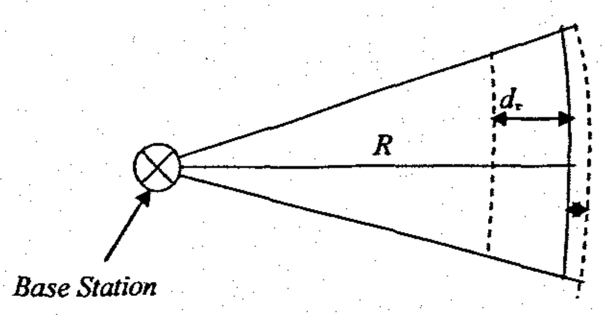

Figure 5.2 Cell sector and HPR region

The HPR is calculated as follows:

$$\% HPR = \frac{d_x}{R} \times 100\ \% \quad \text{.................................. Equation 5.5}$$

Where: HPR = handoff Preparation Region size in percentage

d_x = distance from cell boundary

R = total radius of the cell

5.3 Experiments

The goal of conducting the experiments was to observe and analyse the behaviour of the SAH scheme in managing both new calls and handoff calls. The behavioural analysis experiments were conducted by: *(a)* varying SAH's handoff region under constant speed of MNs, *(b)* varying the speed, keeping the handoff region fixed, and *(c)* varying the load. The first four experiments were conducted exclusively using SAH. They were meant to evaluate the effect of the design criteria (sector approach, and regions). The other sets of experiments were conducted to evaluate the performance of SAH when compared with FHMIPv6. The details of the experiments are tabulated and are also represented graphically. Experiments to measure utilisation were carried out at a cell, chosen as the handoff target cell. This cell accepted both new calls and handoff calls. It is worth noting that, the traffic from all other cells was taken into account.

5.3.1 Effect of HPR region's size variation

(A) Test

The purpose of this experiment was to check how handoff call dropping and new call blocking probabilities are affected by varying the size of the preparation region. For the experiment, the region was varied in only the cell that the mobile node moved from before handing off to another cell. In this case it was SA1_Cell3_BS. The HPR region was varied in this cell. The speed of the mobile nodes was kept constant, at a chosen value, of 40 m/s. SA3_Cell7_BS was made the handoff target cell. The number of calls was set to 4 calls/s. Each mobile node generated one call at any time instant.

(B) Results

The data generated during simulation is presented in a tabular form, in Table 5.1. Figure 5.3 shows a graphical representation of the tabulated data.

Table 5. 1 SAH's HDP and CBP under varying % of HPR (**PHPR:** Percentage (%) of HPR region and **Prob:** probabilities)

PHPR / **Prob**	20	30	40	50	60	70	80
HDP	0.74	0.48	0.27	0.09	0.04	0.06	0.05
CBP	0.79	0.64	0.59	0.58	0.6	0.65	0.68

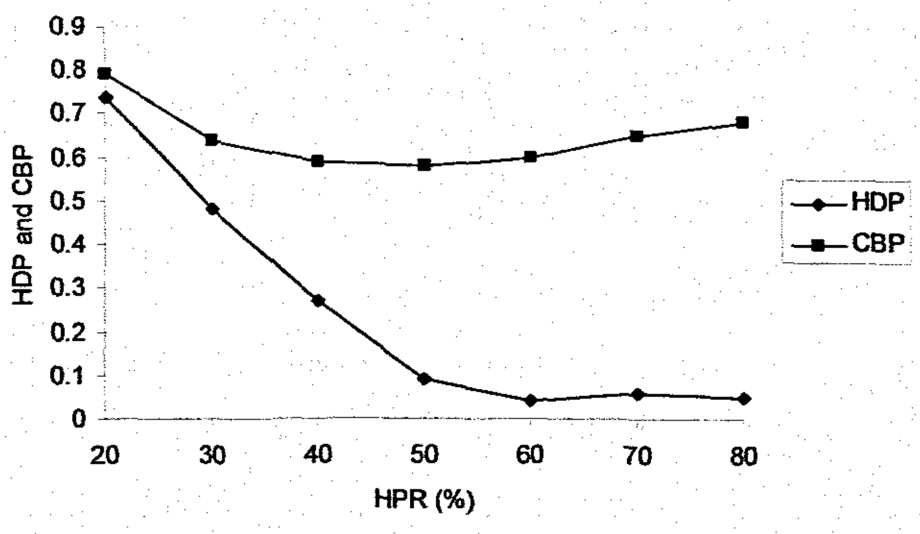

Figure 5.3 SAH New Calls and Handoff Calls

(C) Analysis

From the graph in Figure 5.3, a small HPR, of 20 % produces a high call dropping probability. This is because the mobiles do not get enough time to process the handoff requests since they are moving at high speed of 40 m/s. So there is only a little area (i.e.

shorter distance) available to process all necessary actions. In addition, it could be due to the fact that, the little HPR does not allow more number of retries in requesting resources, in case there is no immediately available free bandwidth in the target cell. As the size of HPR increases, the dropping probability reduces. When the HPR is increased to 50% and above, the dropping probability stays low, less than 0.2. The large size of HPR, 50% to 80% allows mobiles to have more attempts, in requesting for resources but does not reduce the HDP substantially. Above 50 % of HPR, the HDP is approximately zero. The presence of some handoff call drops above the 50 % HPR region may be due to the calls that originate close to the cell boundary. The new call blocking rate stays higher than the handoff call dropping rate, because handoff calls were prioritised over new calls; and the latter had a reserved exclusive share, apart from the common share where both types of calls competed. This characteristic is good since the aim of SAH, just as for any other handoff scheme, is to give more support to handoff calls.

From the results obtained, we could conclude that the optimum size of the HPR region is 50 % of the cell size measured in concentric area. This is because above 50 % of the HPR region, there is no substantial increase in the HDP and CBP, but the delay in using the reserved resources would increase greatly.

5.3.2 Effect of Speed variation

(A) Test

Another design goal was to develop speed-aware handoff preparation and execution decisions. This experiment was conducted to check the behaviour of the SAH scheme when subjected to speed variations of the mobile nodes. In this case, for each experiment, the HPR region was kept fixed at a chosen percentage of 20 %, while the speed was varied from 5 m/s to 40 m/s. The HDP and CBP were observed in cell SA3_Cell7_BS. The number of calls was kept at 4 calls per second.

(B) Results

Table 5.2 shows the data captured during simulation. Figure 5.4, shows graphical representation of the tabulated data.

Table 5. 2 SAH's HDP and CBP under varying Mobile Node's Speed (**Speed:** MNs speed (m/s) and **Prob:** probabilities)

Speed / Prob	5	10	15	20	25	30	35	40
HDP	0	0	0	0.14	0.39	0.57	0.68	0.78
CBP	0.12	0.27	0.38	0.43	0.47	0.48	0.51	0.53

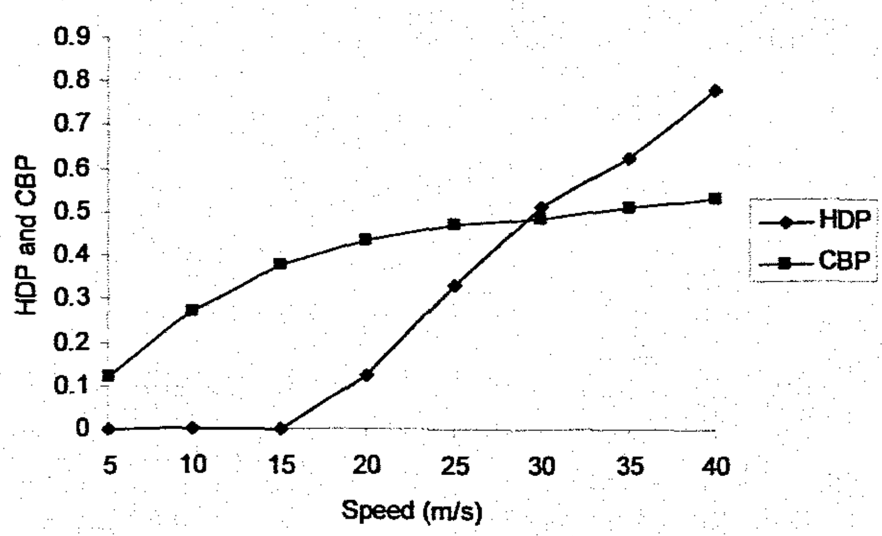

Figure 5.4 SAH speed variation

(C) Analysis

In Figure 5.4, when the speed of mobile nodes is low 5 m/s – 15 m/s, handoff call dropping probability is very low, levelling to zero. The CBP increases since new calls are less prioritised than handoff calls. At this speed, the handoff calls that originate at the boundary are also accommodated. The mobiles are moving slow and the HPR region is big enough, therefore all requesting handoff calls succeed. The unavailability of resources at a certain time does not highly affect MNs' call request, since at this speed they are able to do many retries. As the speed increases, from 15 m/s to 40 m/s, the handoff call dropping probability also increases. This is because at this point, the HDP is affected more by the speed than the HPR size. The dropping probability increases because of the high speed and thus the mobile node could not complete all necessary handoff steps on time. This also shows that the handoff call dropping does not only depend on the instantaneous unavailability of resources, but also on the speed of the MN relative to the handoff area. At the same time, the CBP starts stabilising, increasing slowly. This is because more handoff calls start to be dropped, lowering resource usage; hence new calls take advantage of available resources. From the results, we can conclude that the speed of 15 m/s is the speed at which the system would run optimally.

5.3.3 Effect of Load Variation

(A) Test

This experiment was conducted with the aim of checking the behaviour of the SAH scheme as the load varies. The load was taken to be the number of calls; both new calls and handoff calls that arrived in the cell every period of 1 sec. In this experiment the MNs speed was set to a value of 20 m/s and HPR region was also set to 20 %. The load was varied from 2 calls/s to 18 calls/s. The SA3_Cell7_BS was used as handoff target cell.

(B) Results

The data captured are as shown in Table 5.3. The graphical representation of the data is given in Figure 5.5.

Table 5. 3 SAH's HDP and CBP under varying Load (**Load:** No. of calls/s and **Prob:** probabilities)

Load / Prob	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
HDP	0.01	0.11	0.19	0.24	0.29	0.31	0.33	0.34	0.35	0.4	0.47	0.56	0.68	0.76	0.89
CBP	0.22	0.37	0.51	0.56	0.58	0.59	0.62	0.61	0.63	0.71	0.79	0.88	0.96	1	1

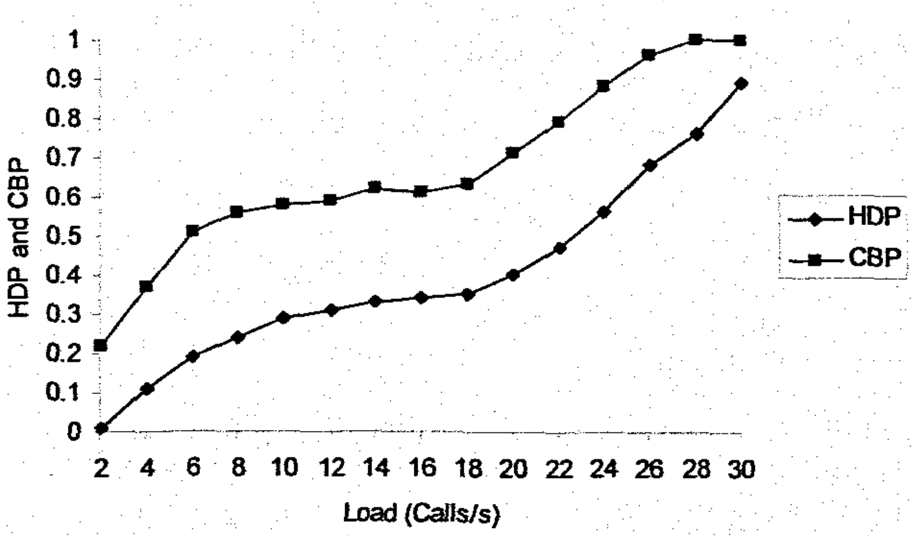

Figure 5.5 SAH load variation

(C) Analysis

Figure 5.5, shows the effect of varying the load in the SAH scheme. As the load is increased from 2 calls/s to 18 calls/s, both the new call blocking and handoff call dropping probabilities increase logarithmically. When the load becomes greater than 20 calls/s, the system is subjected to high stress due to overloading, hence the two probabilities increase steeply. The results obtained show that as the load increases, there is also an increase in both handoff call dropping and new call blocking probabilities.

5.3.4 Speed and HPR size balancing

(A) Test

This experiment was conducted to find out the best possible HPR size that could be recommended for a particular speed of an MN. The call rate was set to 4 calls/s. SA3_Cell7_BS was the handoff target cell. The resources were made readily available, so that any handoff call dropping will not be due to handoff resources problem, but, would be due only to the speed of the mobile node. The major parameters in question were: the speed and size of the HPR. For a particular speed of MN, the HPR was varied until the HDP becomes approximately zero. This was taken to be the point where the values of the two parameters affected the system optimally.

(B) Results

Table 5.4 represents data captured during the simulation run. Figure 5.6 shows a graph of the captured data.

Table 5. 4 SAH's HDP with Speed and HPR balancing (**Speed:** MNs speed (m/s), **PHPR**: percentage (%) of HPR region and **Prob**: handoff dropping probability)

Speed, PHPR / Prob	5	10	15	20	25	30	35	40
	20	20	20	20	30	40	50	60
HDP	0	0	0	0.11	0.1	0.09	0.07	0

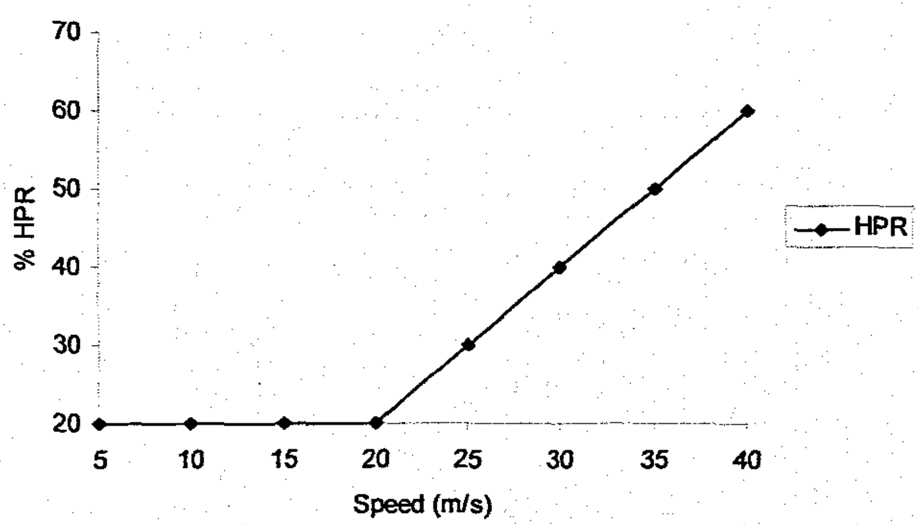

Figure 5.6 Speed and HPR balancing

(C) Analysis

The graph in Figure 5.6 shows an increase in the speed of mobile nodes being counteracted by increase of the HPR. When the mobile nodes move with a speed from 5 m/s to 20 m/s, the HPR region that should be assigned to these MNs for handoff should stay at 20 %. At a speed of 25 m/s, the HPR region becomes small to support mobiles at this speed. Therefore, in order to maintain the same probability of handoff call dropping, the size of HPR needs to be increased to 30 %. When the mobile node's speed increases after 20 m/s there should be corresponding increase in the HPR size recommended to the MNs for handoff, in order to maintain low handoff call dropping probability.

Therefore it can be deduced that, the speed of a MN can be balanced by the size of the HPR in order to maintain a target HDP.

5.3.5 Handoff call dropping probability: Comparison of SAH with FHMIPv6

Test 1: Effect of speed variation on both SAH and FHMIPv6

This experiment was conducted to compare the performance of the SAH scheme and FHMIPv6 scheme in terms of handoff call dropping rate. In this experiment, for each scenario, two cells were monitored, viz: SA1_Cell3_BS and SA3_Cell7_BS. The mobile node's speed was varied from 5 m/s to 40 m/s. The HPR size of SAH was fixed at 20 %.

(A) Results

The results of this experiment are tabulated in Table 5.5. The data captured shows the number of dropped handoff calls. In each experiment, mobiles were subjected to the same speed. Figure 5.7 shows a graph of the captured data.

Table 5. 5 SAH and FHMIPv6 HDP with speed variation (**Speed:** MNs speed (m/s) and **Prob:** handoff dropping probabilities)

Speed / **Prob**	5	10	15	20	25	30	35	40
SAH HDP	0	0	0	0.12	0.28	0.51	0.66	0.78
FHMIPv6 HDP	0	0	0	0.24	0.5	0.73	0.91	0.97

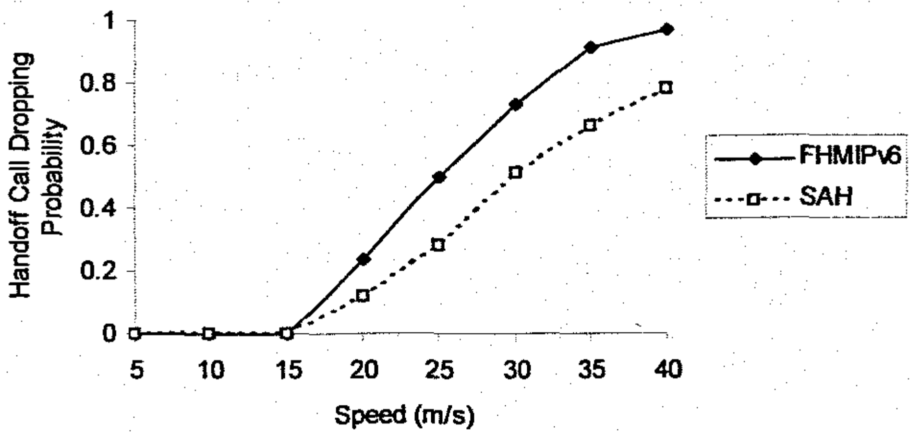

Figure 5.7 Handoff Call Dropping vs Speed

(B) Analysis

At a speed range of 5 m/s to 15 m/s, both schemes are levelled off, the handoff call dropping is zero. This is because when the MNs are moving at low speed, there is enough time for preparing handoff. When the mobile node's speed is above 15 m/s the schemes start experiencing an increase in handoff call dropping probabilities. From the graph we can deduce that as the mobile nodes' speed increases from 15 m/s, there is also an increase in handoff dropping probability. The SAH scheme shows a lower rate of increase of dropping probability with an increase in speed. On the contrary, the handoff dropping probability rate of the FHMIPv6 scheme increases sharply, higher than that of the SAH scheme, with the same increase in speed. This is because with the FHMIPv6 scheme, MNs are left until they are close to the boundary before handoff preparation begins. This presents a very small handoff region for fast moving mobiles, hence the possibility of many handoff calls being dropped. SAH therefore, even when using its smallest size of HPR, out-performs the FHMIPv6 in keeping the handoff call dropping probability low. If

SAH's HPR size is increased to balance the increase in speed of the mobile nodes, this would further lower the handoff call dropping of the SAH scheme. However, since the FHMIPv6 scheme being compared with had no varying HPR, the smallest size of HPR was used for equal opportunity in both schemes.

Test 2: Effect of load variation on both SAH and FHMIPv6

This experiment was conducted to compare the performance of the SAH scheme and FHMIPv6 scheme when load was varied. The load was varied increasingly from 2 calls/s to 18 calls/s. The mobile nodes' speed was set to 20 m/s. The HPR region used was set to 20%.

(A) Results

The results of this experiment are tabulated in Table 5.6. Figure 5.8 shows a corresponding graphical representation of the results.

Table 5. 6 SAH and FHMIPv6 HDP with load variation (**Load:** No. of Calls/s and **Prob:** handoff dropping probabilities)

Load / **Prob**	2	4	6	8	10	12	14	16	18
SAH HDP	0.01	0.11	0.19	0.24	0.29	0.31	0.33	0.34	0.35
FHMIPv6 HDP	0.1	0.18	0.27	0.33	0.39	0.41	0.43	0.43	0.44

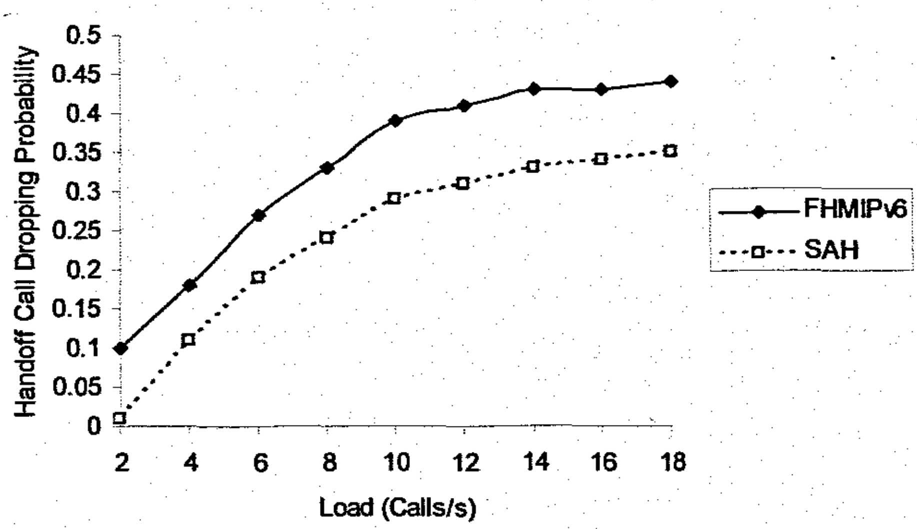

Figure 5.8 SAH vs FHMIPv6 under Load Variation

(B) Analysis

The graph in Figure 5.8 shows that, as the load increases the handoff dropping probability also increases in both schemes. The SAH scheme always has a lower handoff call dropping probability than the FHMIPv6 scheme for corresponding values of load. This is because in the SAH scheme, the handoff preparation begins early as soon as the MN enters the HPR region; and also, prediction facilitates the discovery of the new access router quickly. It does not wait for a specific point or some indications from the link layer.

5.3.6 New call blocking probability: Comparison of SAH with FHMIPv6

Test 1: Effect of speed variation on SAH and FHMIPv6 schemes

This test was conducted at the level of new call blocking probability in the SAH scheme and FHMIPv6 scheme. While concentrating on new calls, the handoff calls were also admitted in the system, thus keeping the system balanced. The mobile nodes' speed was varied from 5 m/s to 40 m/s. The HPR in SAH was set to 20 %.

(A) Results

Table 5.7 present results captured during the simulation run, in tabular form. The data shows the number of blocked new calls in each scheme. Figure 5.9 represents the captured data in graphical form.

Table 5. 7 SAH and FHMIPv6 CBP with speed variation (**Speed:** MNs Speed (m/s) and **Prob:** new call blocking probabilities)

Speed / **Prob**	5	10	15	20	25	30	35	40
SAH CBP	0.12	0.27	0.38	0.43	0.47	0.48	0.51	0.53
FHMIPv6 CBP	0.08	0.24	0.33	0.38	0.37	0.36	0.34	0.32

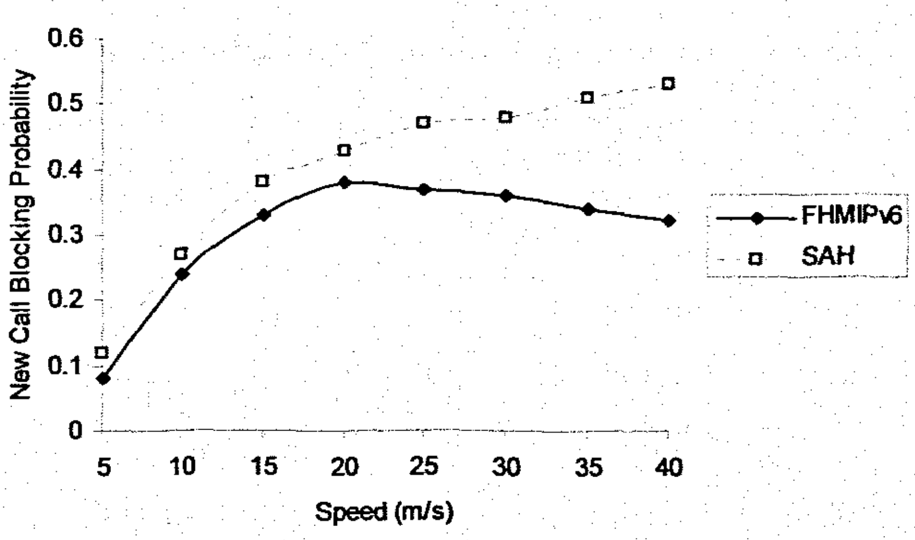

Figure 5.9 New Call Blocking vs Speed

(B) Analysis

The graph in Figure 5.9 shows that, when the mobile nodes enter cell SA3_Cell7_BS moving at a low speed, 5 m/s to 15 m/s, the new call blocking probability in this cell increases in both schemes. The SAH scheme in this case has a higher call blocking probability when compared to FHMIPv6 because, SAH's handing-off calls persist long enough when requesting resources, thus hindering new calls from gaining access.

When the speed becomes high (i.e., from 20 m/s and above), the new call blocking probability for FHMIPv6 decreases. This is simply because the increase in speed increases the handoff call dropping rate, thus leaving more resources unused. The new calls then take advantage of these unused resources. However, the increase in speed does not seriously affect the new call blocking probability in SAH. Therefore, FHMIPv6 is regarded as being better than SAH in terms new call blocking probability.

Test 2: Effect of load variation

The purpose of this experiment was to evaluate the change of new call blocking probability with a varying load parameter. The load was varied from 2 calls/s to 18 calls/s. The MNs moved at a speed of 20 m/s. The HPR size for SAH was 20 %.

(A) Results

Table 5.8 is a representation of the data collected during the simulation. Figure 5.10 shows a corresponding graphical representation of the obtained results.

Table 5.8 SAH and FHMIPv6 CBP with load variation (**Load:** No. of Calls/s and **Prob:** new call blocking probabilities)

Load \ **Prob**	2	4	6	8	10	12	14	16	18
SAH CBP	0.22	0.37	0.51	0.56	0.58	0.59	0.62	0.61	0.63
FHMIPv6 CBP	0.16	0.24	0.31	0.31	0.33	0.36	0.37	0.43	0.41

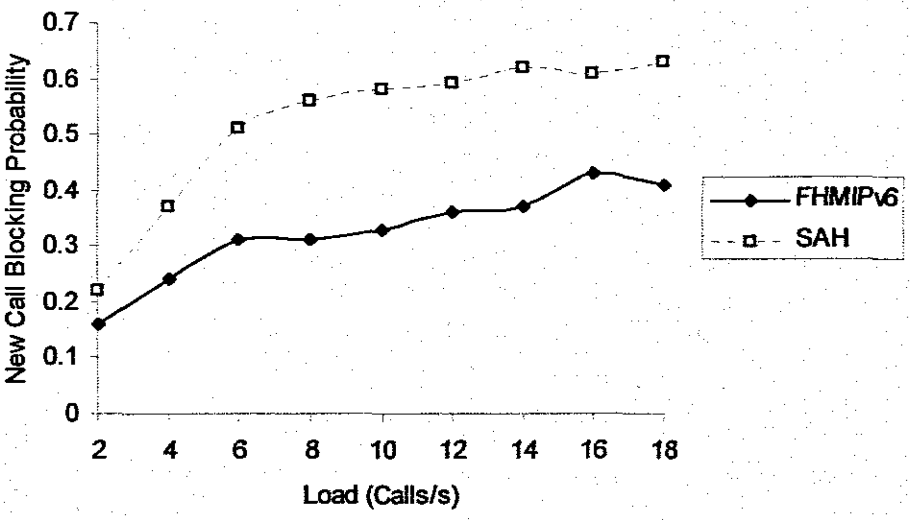

Figure 5.10 SAH vs FHMIPv6: Load Variation

(B) Analysis

The results in Figure 5.10 show that as the load is increased, the call dropping rate also increases in both schemes. Both schemes start with a blocking probability higher than zero at low load, 2 calls/s. As the load is increased to 5 call/s and higher, both schemes start levelling off, but SAH has higher new call blocking probability. In this experiment we see once again that SAH's blocking probability remains higher than that of FHMIPv6. Therefore it can be deduced that, FHMIPv6 out performs SAH in terms of new call blocking probability with varying load.

5.3.7 Utilisation: Comparison of SAH with FHMIPv6

Test 1: Effect of speed variation

This experiment was conducted to compare the utilisation factor in both schemes (SAH and FHMIPv6) as the speed was varied. The speed of the mobiles was varied increasingly from 5 m/s to 40 m/s. The statistical data was collected from cell SA3_Cell7_BS. This cell was also made the handoff target cell; receiving both new calls originating in it, and handoff calls from other cells.

(A) Results

Table 5.9 presents the results captured during the simulation run, in a tabular form. Figure 5.11 represents the captured data in graphical form.

Table 5. 9 SAH and FHMIPv6 Utilisation with speed variation (**Speed:** MNs Speed (m/s) and **Util:** Utilisation)

Speed / **Util**	5	10	15	20	25	30	35	40
SAH	90	86	86	88	87	82	78	78
FHMIPv6	92	88	89	79	76	69	56	48

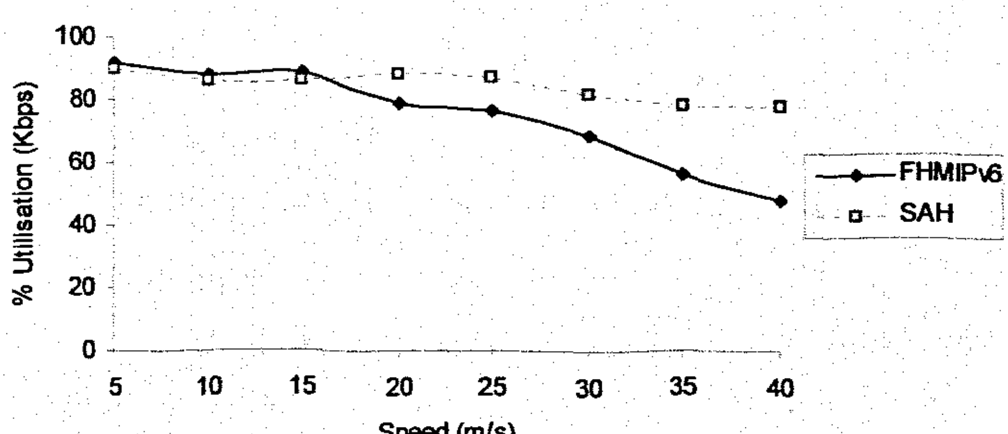

Figure 5.11 Utilisation vs Speed

(B) Analysis

In the graph of Figure 5.11, it can be observed that at a low speed range (5 m/s to 15 m/s) both schemes have high utilisation, of above 80 %. At this speed range, both schemes can effectively support handoff calls (as observed in previous graphs). When the speed is increased above 15 m/s, the utilisation in both schemes begins to decrease. However, SAH's utilisation remains higher than FHMIPv6. This is simply because SAH is less affected by increase in speed; and so it continues supporting both types of calls effectively. FHMIPv6, on the other hand, cannot not effectively support high speeding mobiles. Therefore the handoff calls of mobile nodes are subjected to higher call dropping, and thus leading to under-utilisation.

Test 2: Effect of load variation

This experiment was conducted to evaluate how the utilisation of both schemes, varied with the load parameter. The speed of the mobile nodes was set to 20 m/s. The HPR size from SAH was set to 20 %.

(A) Results

Table 5.10 shows the data collected during the simulation run. Figure 5.12 represents the corresponding data in a graphical form.

Table 5. 10 SAH and FHMIPv6 Utilisation with load variation (**Load:** No. of Calls/s and **Util:** Utilisation)

Load / Util	2	4	6	8	10	12	14	16	18
SAH	42	66	72	72	74	67	63	65	67
FHMIPv6	49	71	83	85	89	87	88	82	86

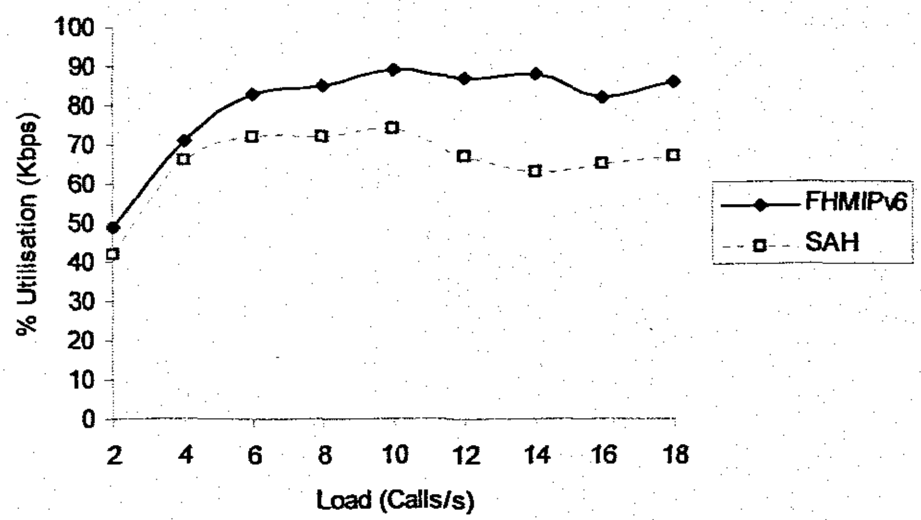

Figure 5.12 Load vs Utilisation

(B) Analysis

The graph in Figure 5.12 shows that both schemes tend to increase in utilisation as load increases from 2 call/s to 6 calls/s. SAH in this case starts with a low utilisation and remains lower than FHMIPv6 throughout the experiment. In principle, SAH reserves resources early, and may not use the resources immediately, so much resource is blocked as reserved. This reduces the utilisation. FHMIPv6 reserves resources and makes use of the reserved resources immediately. Therefore SAH, performs lower than FHMIPv6 in terms of utilisation as load increases. As the load further increases from 6 call/s to 10 call/s, the utilisation remains almost constant. Despite the load increase, both schemes remain relatively at constant levels, but with SAH always at a lower Utilisation.

In summary, on the one hand, the SAH scheme performs better than the FHMIPv6 scheme in terms of handoff call dropping probability when the speed or the load increases. On the other hand, the FHMIPv6 scheme is better than the SAH scheme in new call blocking. The SAH scheme, in addition performs better in terms of utilisation when there are high speed mobile nodes.

Chapter Six

CONCLUSION

6.1 Introduction

Network layer handoff delays and service disruptions experienced by real-time applications, have led to the development of fast handover protocols that anticipate L3 handover before its actual occurrence. However, these protocols lack good support for high speed mobile nodes. To provide fast handover and proper support for MN's speed variations, a Sector Aware Handoff scheme (SAH) has been proposed. SAH offers better means of predicting L3 handover, and low handoff call dropping probability.

This chapter reviews the SAH handoff scheme. It also reviews the achievements of the research work. Reasonable critiques of the work and some suggestions on how it can be extended in future have also been included. The chapter concludes by outlining some directions of future work.

6.2 Summary

In order to develop SAH, relevant design characteristics were identified. These are: (i) Early anticipation of L3 handover – procedure of predicting the network layer handoff before its occurrence; (ii) Mobile Node's speed-aware handoff preparation and execution decisions – handoff triggering decision that uses mobile node's active mobility profile in real-time; and (iii) Enhanced prioritisation in handoff queue (i.e., strategic access to handoff resources) – the strategic dynamic handoff queue and reshuffling technique. In addition to the most basic employed prioritisation technique, (i.e., MN's position from cell boundary), the type of application being transmitted and the mobile node's real-time speed are also taken into account. In SAH these characteristics were improved to meet real-time applications' QoS requirements, and to effectively support fast moving mobiles.

The SAH scheme employs a hierarchical architecture. The components of this architecture are: domains, controlled by a domain agent; and subnets, controlled by a subnet agent. Each subnet is made up of seven cells. The model uses sectored cells that have virtual handoff regions, for handoff preparation and handoff execution (HPR and HER regions), hence handoff in our system is divided into two phases, preparation phase, and execution phase. There are six sectors in a cell. Cell sectors are marked into three virtual regions, the non-handoff region (NHR), handoff preparation region (HPR) and handoff execution region (HER). HPR is an area between non-handoff region (NHR) and the HER. HPR is adjustable, i.e., can be made large or small, depending on mobile node's speed. HER is an area close to the boundary of the cell. NHR is an area close to the base station. These regions assist in terms of where and when a particular handoff phase should be carried out, depending on a mobile node's speed.

When a mobile node enters HPR region, its real-time mobility profile is analyzed to check whether it is necessary to initiate handoff preparation. If so, then the preparation phase commences, beginning by invoking a next-cell prediction process. The next-subnet-cell prediction is facilitated by a next-cell prediction scheme proposed by (Nkambule et al, 2004). Thus prediction is done anywhere within the cell sectors when deemed so. The activities in the preparation phase include: next-subnet cell prediction, handoff resource acquisition, etc. This phase finishes by indicating the status "Ready", before the execution phase takes place. When the mobile node enters the HER region, the preparation status is checked. If the status is ready, execution takes place; otherwise, the call is dropped. The main task in the execution phase is switching between base stations.

The model offers better and realistic L3 prediction. It uses MN's real-time mobility parameters that may be obtained using a positioning technique, such as GPS. Hence, this offers a better approach for early anticipation of L3 handover.

Extensive simulation was performed using parameters such as: (i) mobile nodes' speed, (ii) size of HPR region, and (iii) Load. The performance metrics used were: handoff dropping probability (HDP), new call blocking probability (CBP), and utilisation. Further, a comparison of the SAH and FHMIPv6 scheme was made.

Results achieved show that SAH performs better in supporting speed variations in MNs, and in minimising handoff call dropping probability. It was further observed that SAH performs better when compared to FHMIPv6 in terms of call dropping probability, and utilisation, especially when there are many high speed mobile nodes present. FHMIPv6 on the other hand, shows a better performance over SAH with respect to new call blocking probability.

The SAH scheme however, has some limitations. Theoretically, it has high signalling, in terms of control and request messages, especially during the resource acquisition phase; in cases where resources are not immediately available. SAH has a disadvantage that the handoff reservations that have been made may take a long time before they are used.

A limitation of the simulator was that the total delay in L2 handoff was assumed to be a specific value in milliseconds. However, in reality this time can vary depending on the association delays, frame transmission delays, etc.

Nevertheless, the concept of SAH can be the basis for providing fast handoffs for high speed mobiles in the Wireless Internet. It has a dynamically adjustable handoff region depending on a mobile node's speed. It also offers better treatment for L3 handovers by modularising the handoff process into separate phases i.e., preparation, and execution phases. It performs early anticipation of L3 handover through prediction, using real-time mobility parameters of the mobile node over a confined cell area; the sector.

6.3 Future work

Simulation results obtained show that SAH is a better candidate in providing fast L3 handoff. In order to improve the SAH strategy, additional work could be done by: (i) incorporating a more sophisticated call admission criterion, (ii) developing L2 handoff strategies that can be able to predict forthcoming L2 handover as early as possible, without highly relying on signal characteristics; instead, by using real-time mobility parameters of MNs and some real-time stored data. This can further increase the performance of SAH, since handoff delays posed by L2 also contribute to L3 handoff delays. (iii) using a Poisson traffic model in addition to the currently used random model.

This research concentrated on the call level aspect of network layer handoff, with the use of call level performance metrics such as, handoff call dropping probabilities, new call blocking probabilities, and system utilisation. The performance metrics investigated were also applicable to call level in handoff. However, there is an understanding that the performance metrics used in the study, only form a small part of the performance metrics that can be investigated in a handoff domain. One of the important tasks for future work in SAH, may be to investigate its performance with packet level handoff. The performance metrics to be investigated in this level could include (but not limited to): handoff delay, latency, packet loss, and buffering, as these are some of the most important QoS characteristics with respect to real-time applications.

Appendix

A-1 Class Diagram and Class Description

Our simulation has twelve classes, shown in Figure A-1. Classes are objects having properties and methods. They carry out specific functions and these functions are coordinated to perform a specific task. All the classes work in a coordinated manner resulting in a unit system.

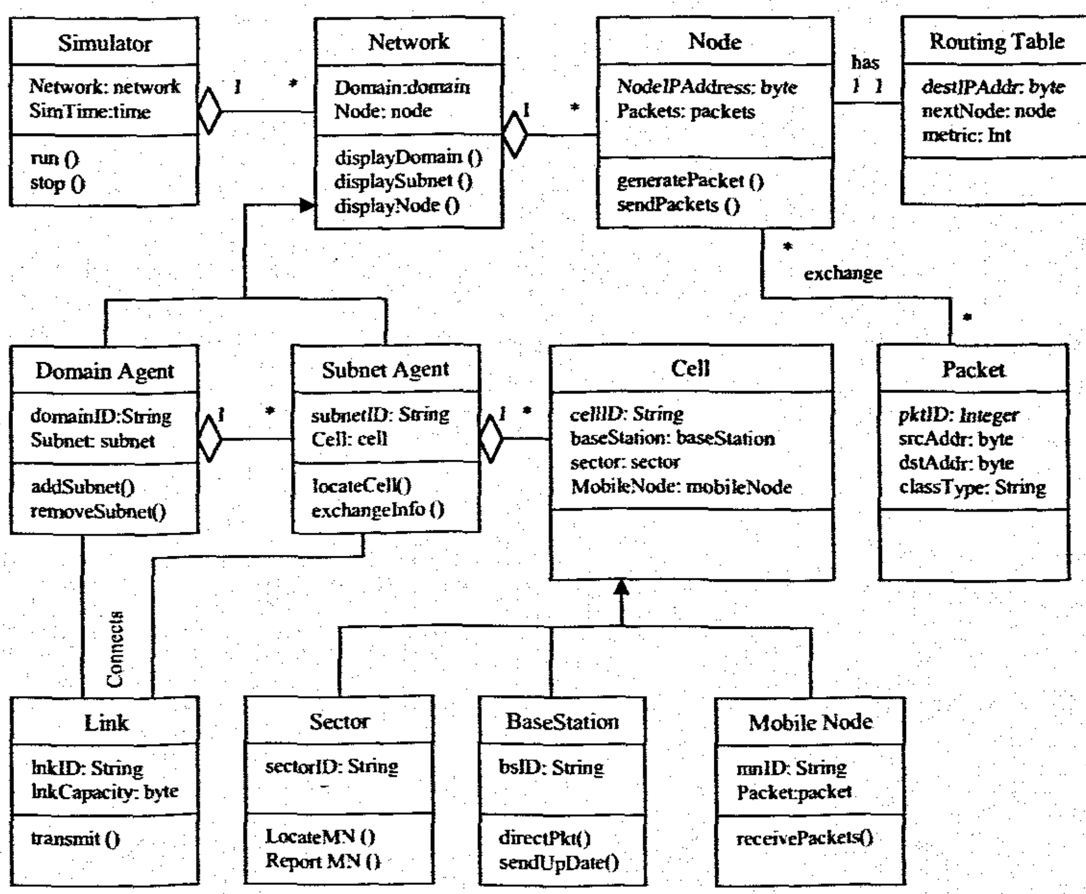

Figure A- 3 1 SAHS Class Diagram

Class Description:

Simulator: This class is a container class that holds all simulation classes. It provides interface to users to set parameters when running the simulation.

Network: Network class contains all node classes. Those nodes act as components building up a network.

Node: This is the base class with generic methods and function for network entities; Domain Agents, Subnet Agent, Base Stations, and mobile nodes.

Routing Table: This class is a component of a node class. It contains entries about Mobile nodes' subnet residence, represented by their destination IP addresses.

Domain Agent: This class extends the node class. It controls several subnets. It has statistical table about the residence of mobile nodes in subnets (i.e., keeps MN's entries per subnet.). It redirects packets in accordance to mobile nodes' subnet location.

Subnet Agent : This class also extends the node class. It controls a group of Base Stations, and provides a finer location tracking of Mobile nodes. It receives reports about MN's behaviour through Base Stations and also exchanges information with its neighbours.

Cell: This class represents a service area for mobile nodes. It comprises of Base Station, and sectors.

Packet: This class represents a message container. It has source address and destination addresses.

Link: This class represents a channel that connects nodes from access network to core network i.e., subnets and domains. It transmits packets between nodes.

Sector: This class represents an area in a cell which provides more refined location of mobile nodes.

Base Station: This class represents a node. It delivers packets to registered mobile nodes residing in its coverage area. It also monitors the behaviour of mobile nodes i.e. where the mobile nodes move to.

Mobile Node: This class also extends the node class. It creates a mobile node object with attributes (speed, bandwidth requirements, etc) and methods (move, stop, accelerate, etc)

A-2 Simulator User Interface

Figure A-2 shows user interface of the simulator in a non-running mode.

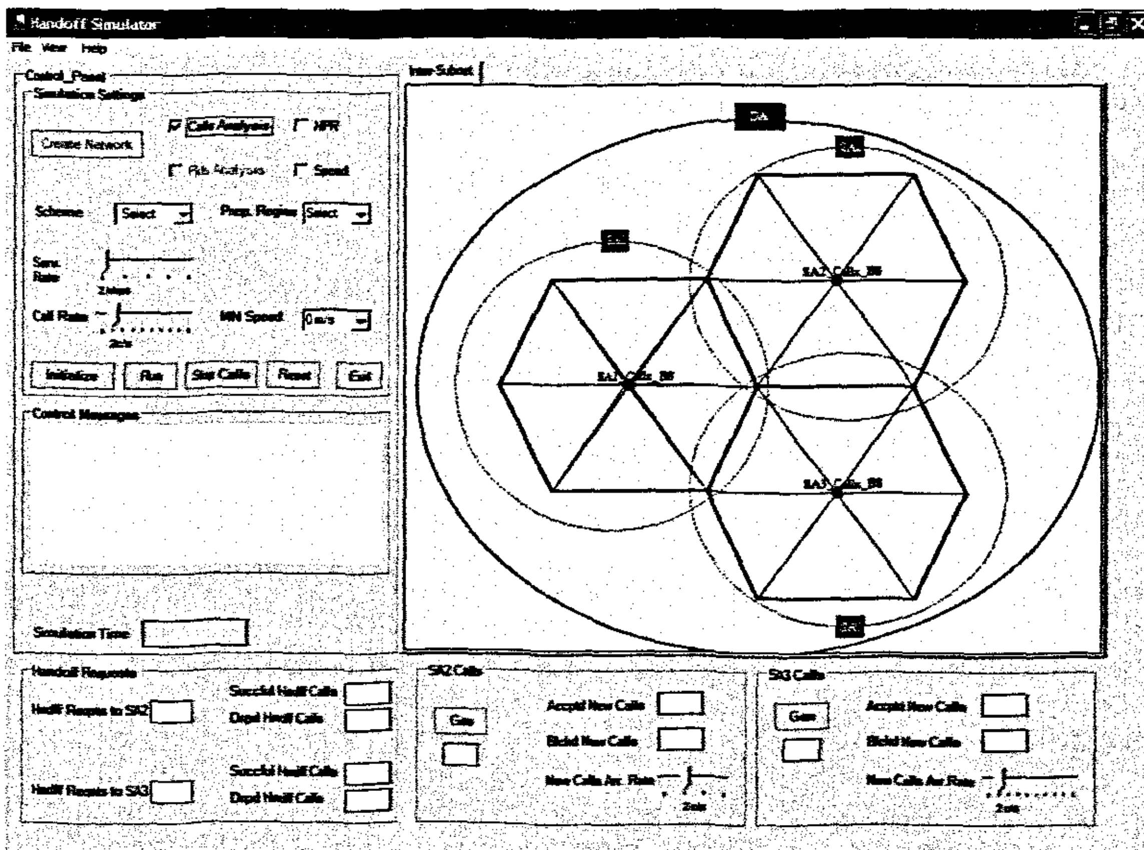

Figure A- 2 Simulation user interface

From the interface in Figure A-2, the simulator shows different panels i.e. a control panel where simulation settings are done, with respect to experiment being conducted. Control messages panel, that displays requests and reply messages being transmitted by the network and nodes. Statistical panels, bottom panels that show information such number of successful and dropped handoff calls, number of successful and blocked new calls etc. This information can be shown for per subnet.

The create network button draws the network entities, DA, SAs and BSs with cells as objects. When the network has been created, the user checks the appropriate checkbox HPR, Speed. The scheme combo box allows the user to select a scheme wanting to do simulation with. There are two schemes, SAH and FHMIPv6 loaded in the combo box. Depending on the scheme selected, some properties e.g. HPR checkbox, Prep region combo box are disabled or remain enabled. For an example, when FHMIPv6 scheme is selected, HPR is disabled, because FHMIPv6 scheme has no HPR.

The other settings include service rate, call rate, MN speed. The initialise button is clicked when all desired settings have been set, then on initialisation; the system configures itself for relative to settings made. The run button starts the simulation. Mobile nodes are created, calls are made. The calls are generated by host outside the domain, and enter through domain to mobiles moving around subnet cells. The stop calls button terminates generation of calls by mobile nodes. This same button when clicked from stop calls, it turns to stop sim button, to stop the simulation. Reset button resets the settings to default values. The other controls are for SA2, and SA3 calls statistics.

Figure A -3, shows a captured screen short of the simulator in a running mode.

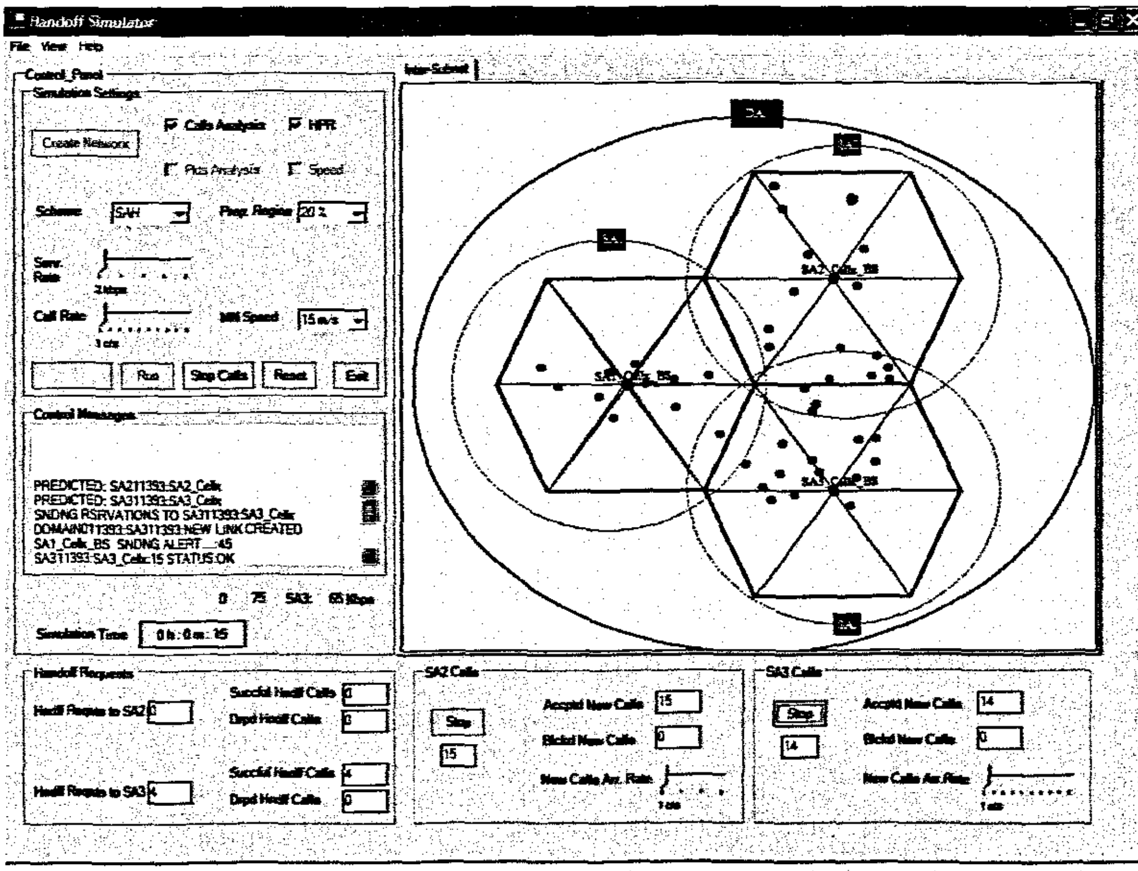

Figure A- 3 Simulator in running mode

References

Abdulova, V. and Aybay, I. (2006). "Handoff Prioritizing Schemes in Cellular Wireless Networks", *Computer Networks*, 2006 International Symposium on OPAC, June, pg (s). 1-6.

Choi, C.H. Kim M. Kim T. J. and Kim S. J. (2000). "Adaptive Bandwidth Reservation Mechanism Using Mobility Probability in Mobile Multimedia Computing Environment", Proceedings of the 25th Annual IEEE Conference on Local Computer Networks (LCN.00), p. 76.

Chou, C. and Shin K.G. (2005). "An Enhanced Inter-Access Point Protocol for Uniform Intra and Intersubnet Handoffs", IEEE Transactions on Mobile Computing, July, pp. 321-334.

Chung, S. and Chien, S. (2005). "Analysis of multiservice cellular networks with asymmetrical traffic and handoff queue", Journal of Computer Communications, vol. 28 (2005), pp 864 -879. Available online, December 2004.

Dommety, G. Yegin, A. Perkins, C. Tsirtsis, G. El-Malki, K. and Khalil, M. (2002). "Fast handovers for mobile IPv6," Internet draft, draft-ietf-mobileip-fastmipv6-04.txt, March. Last accessed: 27 December 2006.

Duong, H. Dadej, A. and Gordon, S. (2005). "Proactive Context Transfer and Forced Handover in IEEE 802.11 Wireless LAN Based Access Networks", ACM SIGMOBILE Mobile Computing and Communications Review, Volume 9, Issue 3 (July 2005), pp 32 – 44.

Feng, F. and Reeves, D.S. (2004). "Explicit Proactive Handoff with Motion Prediction for Mobile IP", IEEE Wireless Communications and Networking Conference, 2004. WCNC. 2004, Volume: 2, pp 855- 860.

First National Bank, (2006), https://www.fnb.co.za/siteinfo/information/onlineBank.html, last accessed: 30 November, 2006.

Forouzan, B.A. (2004). "Data Communications and Networking", third edition, New York, McGraw-Hill, publishers.

Fournogerakis, P. Kyriazakos S. and Karetsos, G. (2001). "Enhanced Handover Performance in Cellular Systems based on Position Location of Mobile Terminals", Proceedings of the IST Mobile Summit, 2001.

Helmy, A. Jaseemuddin, M. and Bhaskara, G. (2002). "Efficient Micro-Mobility using Intra-domain Multicast-based Mechanisms (M&M)", ACM SIGCOMM Computer Communication Review, Vo. 32, Issue 5 (November 2002), pp: 61 – 72.

Hernandez, E. and Helal, S. (2004). "Predictive Mobile IP for Rapid Mobility", Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks (LCN'04) , pp 684- 6896, November.

Huang, N. and Chen, W. (2003). "RSVP Extensions for Real-time Services in Hierarchical Mobile IPv6", Journal of Mobile Networks and Applications vol.8, pp 625 – 634, Kluwer Academic publishers.

Iera, A. Molinaro, A. and Marano, S. (2002). "Handoff Management With Mobility Estimation in Hierarchical Systems", IEEE transactions on vehicular technology, vol. 51, no. 5, pp 915- 934, September.

Ivanov, K. Spring, G. (1995). " Mobile Speed Sensitive Handover in a Mixed Cell Environment", IEEE 45^{th} Vehicular Technology Conference, vol. 2, pp 892 – 896, July.

Jung, H. Kim, E. Yi, J. and Lee, H. (2005) [a]. "A Scheme for Supporting Fast Handover in Hierarchical Mobile IPv6 Networks", ETRI Journal, Volume 27, Number 6, December.

Jung, H. Kim, E. Yi, J. and Lee, H. (2005) [b]. "Fast Handover for in Hierarchical MIPv6 (F-HMIPv6)", IETF draft, draft-jung-mobopts-fhmipv6-00.txt, April. Last accessed: 30 November, 2006.

Kempf, J. Funato, D. El-Malki, K. Gwon, Y. Petterson, M. (2001). "Supported Optimized Handover for IP Mobility – Requirements for Underlying Systems", Internet engineering Task Force Team (IETF) draft-manyfolks-12-mobilereq-01.txt, Nov. Last accessed: 30 November, 2006.

Kim, H. and Jung, J. (2001). "A Mobility Prediction Handover Algorithm for Effective Channel Assignment in Wireless ATM," IEEE GLOBECOM, vol. 6, 2001, pp. 3673-3680.

Langar, R. Bouabdallah, N. and Tohme, S. (2006). "Handoff Support for Mobility in Future Wireless MPLS Networks: a Proposal and Analysis", IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006, Vol. 1, pp: 556- 561, April.

Lee, C.Y. and Kang, H.G. (2000). "Performance of the Double Tunneling with Handoff Prediction in Mobile IP Networks", Proceedings of Wireless 2000: The 12th international conference on wireless communications, 2000, pp. 193-202, July.

Lin, J. C. and Rangarajan, S. R. (2006). "LIHP: A Low Latency Layer-3 Handoff Scheme for 802.11Wireless Networks", Proceedings of IEEE 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'06), pp. 401-409.

Lo, S. Lee, G. Chen, W. and Liu, J. (2004). "Architecture for Mobility and QoS Support in All-IP Wireless Networks", IEEE Journal on Selected Areas in Communications, vol. 22, No. 4, May.

Lopez, A. Manner, J. Mihailovic, A. Velayos, H. Hepworth, E. and Khouaja, Y. (2001). "A Study on Provision for IP-based radio Access Networks", Proceedings of the Tyrrhenian International Workshop on Digital Communications, Taormina, Italy, September.

Lu, L. Wu, J.C. and Chen, W. (2004). "The Study of handoff prediction schemes for Resource Reservation in Mobile Multimedia Wireless Networks", Proceedings of the 18th International Conference on Advanced Information Networking and Application (AINA'04), March.

Nkambule, M.W. Ojong, G.E. and Adigun, M.O. (2004). "Predicting mobility on the Wireless Internet Using Sectored-cell Approach", Proceedings of SATNAC 2004, September.

Pack, S. and Choi, Y. (2002). "Pre-Authenticated fast handoff in a public wireless LAN based on IEEE 802.1x model", Proceedings of the IFIP TC6/WG6.8 Working Conference on Personal Wireless Communications, pp. 175 - 182, October.

Pack, S. Jung, H. Kwon, T. and Choi, Y. (2005). "SNC: A Selective Neighbor Caching Scheme for Fast Handoff in IEEE 802.11 Wireless Networks" ACM SIGMOBILE Mobile Computing and Communications Review, Vol. 9, Issue 4 (October 2005).

Park, B. Lee, S. and Latchman, H. (2006). "A Fast Neighbor Discovery and DAD Scheme for Fast Handover in Mobile IPv6 Networks", Proceedings of the International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06), April.

Perkins, C. (1996). "IP Mobility Support", Network Working Group. RFC 2002, October 1996.

Perkins, C. (2002). "IP Mobility Support for IPV4", RFC 3344, August, 2002.

Salkintzis, A.K. (2004). "Mobile Internet, Enabling Technologies and Services", CRC Press, New York.

Samprakou, I. Bouras, C. Karoubalis, T. (2005). "Fast IP Handoff Support for VoIP and Multimedia Applications in 802.11 WLANs", Sixth IEEE International Symposium, on a World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005., pp: 332-337, June.

Saraswady, D. and Shanmugavel, S. (2004). "Seamless Handoff Architecture: A Micromobility Management for Next Generation Wireless IP Networks", Information Technology Journal vol.3 (3): pp. 319-326.

Shim, E. and Gitlin, R. (2001). "Neighbor Casting: A Fast Handoff Mechanism in Wireless IP Using Neighbouring Foreign Agent Information", *New York Metro Area Networking Workshop*, 2001.

Shim, E. Wei, H. Chang, Y. and Gitlin, R. D. (2002). "Low Latency Handoff for Wireless IP QOS with NeighborCasting", IEEE International Conference on Communications, 2002. ICC 2002., Vol.5, pp.3245- 3249.

Shyy, D. (2004). "Requirements for Different Handoff Scenarios", IEEE 802.21 WG, doc.: IEEE 802.21-04/xxxr0, March.

Soliman, H. Castelluccia, C. El-Malki, K. and Bellier, L. (2002). "Hierarchical Mobile IPv6 mobility management," Internet draft, draft-ietf-mobileip-hmipv6-07.txt, October, Last accessed: 30 November 2006.

Sung, C. W. and Wong, W. S. (1994). "User Speed Estimation and Dynamic Channel Allocation in Hierarchical Cellular System", Proceedings of IEEE 44th Vehicular. Technology Conference, vol. 1, 1994, pp. 91-95.

Tan, C. L. Pink, S. and Lye, K. M. (1999). "A Fast Handoff Scheme for Wireless Networks," 2nd ACM International Workshop on Wireless Mobile Multimedia, Washington, USA, August.

Trossen, D. Krishnamurthi, G. Chaskar, H. Kempf, J. (2001). "Issues in Candidate Access Router Discovery for Seamless IP Handoffs", IETF draft, draft-ietf-seamoby-cardiscovery- issues-00.txt, July. Last accessed: 30 November 2006.

Van den Wijngaert, N. and Blondia, C. (2002), "A Predictive low latency handover scheme for mobile IP", Proceedings of OPNETWORK, 2002. August.

Xu, Z. and Ye, Z. Kirshnamurthy, S.V. Tripathi, S.K. and Molle, M. (2002). "A new adaptive channel reservation scheme for handoff calls in wireless cellular networks", LNCS, vol. 2345, pp. 672–684, 2002, ISBN:3-540-43709-6.

Yang, Z. Wang, Y. Zhao, D. He, J. Fuy, X. (2005). "Fast seamless handover scheme and cost performance optimization for ping-pong type of movement", IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. Vol. 4, pp: 2131- 2135.