QUALITY OF SERVICE PROVISIONING IN MOBILE WIRELESS NETWORK

MAFIKA WILLIAM NKAMBULE

A dissertation submitted to the Faculty of Science and Agriculture in Fulfilment of the requirements for the degree

MASTERS OF SCIENCE

In

COMPUTER SCIENCE

Department of Computer Science

University of Zululand

Promoter: Prof. M.O. Adigun

Co-promoter: Mr. G. Ojong

Department of Computer Science

2004

DECLARATION

This dissertation represents research work carried out by the author and has not been submitted in any form to another University for a degree. All the sources used in the dissertation have been duly acknowledged.

Signature: ______________________

DEDICATION

There are those people in our lives who miraculously appear and affect us in ways for which we are ever grateful. They bestow upon us their knowledge and experience and send us on our way. One such man came into my life three years ago in the form of my academic advisor. He nurtured, encouraged, criticized and comforted me. He has now "sent me on my way" – on the new representation. There has been, and will always be, a piece of him in every work I produce. I owe my career to Jabu Bhekizwe Dlodlo. I will miss working with him very much.

ACKNOWLEDGMENTS

First and foremost, I would like to express my heartfelt gratitude to my supervisors, George Ojong, who has shown great care and support; and Professor Mattew Adigun, for his guidance and support as well. He is a great role model for his students. He stimulates critical thinking, gives insightful comments, and is always willing to share his invaluable experiences, which will benefit our careers for the years to come. I want to take this opportunity to thank my fellow colleagues, Themba Khumalo, Philile Biyela, and Hlengiwe (Pinky) Kunene as well as our other research group members Sihle, Sarah, Mathalenta, Mzila and Zandile. I would like to thank them for their friendship, emotional support, numerous intellectual discussions, and occasionally brewing delicious coffee for each other. May our bonds remain strong after we head separate ways to pursue our own professional goals.

I am also grateful to Dr. S.S. Xulu for his friendly help. I would also like to thank Telkom, for funding my studies at Zululand University under their Center of Excellence Programme. Most importantly, I would like to express my grateful appreciation to my parents, brother, and sisters, who have always believed in me, and to my lovely wife Londiwe Mkhwanazi, for her love and care. I cannot imagine going through graduate life without their love and support. Last but not least, my friends S'the and Mkhize who were always there to advice and assist me whenever I needed any help.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

One major challenge in the design of wireless mobile networks is the provision of quality of service (QoS) guarantees. Mobile users would require resources to be available in the cells they would move into in order to maintain the same QoS as in the current cell. However, reservation in all neighbouring cells or based on the user's mobility history results in over-reservation of resources, leading to schemes that are very inefficient and unreliable. This dissertation presents a next-cell prediction scheme (NCPS) that uses a sectored-cell approach to predict a next cell, a mobile that will move into with an open connection. The NCPS was simulated and its performance was evaluated under different speeds of the mobile users. The behaviour of the scheme was also observed under different sizes of the critical region. The results of the scheme were also compared to the results of a representative history-based scheme. It was found that the NCPS performs better than the history-based schemes. The simulation results show that prediction accuracy can be achieved with a considerable decrease in the total tracking area of the network.

CHAPTER ONE

INTRODUCTION

1.1 Overview

The demand for mobile wireless networks has been steadily increasing in the past several years. This demand has pushed the networks in question up to and beyond their handling capacity. New network designs and protocols are needed to support the growing demand for network resources and the upcoming need for quality of service (QoS) guarantees to support multimedia traffic. Thus one major challenge in the design of wireless mobile networks is the provisioning of QoS guarantees that applications demand under this diverse networking infrastructure. We believe that it is necessary to use prediction of a cell where resources can be reserved to deliver these QoS guarantees to applications. However, reservation in all neighbouring cell or in the entire service region [1] can be extremely aggressive, and can result in schemes that are very inefficient and unreliable. To overcome this, the entire radio cell can be partitioned into small areas, called sectors. If a cell can be partitioned such that each sector is adjacent to only one neighbouring cell, we can therefore associate the target cell to each adjacent sector of an active cell. Thus the reservation procedure can be limited to the target cells only. A target cell is assumed to be a cell a mobile user is likely to visit.

Recently, a number of schemes that apply user mobility prediction to various aspects of mobility management have been reported in the literature. It has been shown through modelling and simulation that the use of mobility prediction is effective to enhance the performance of resource reservation [2], [3], [4], [5], [6], [7], handover management [8], [9], [10], and location management [8], [11], [12] schemes. Furthermore, it has also been shown that movement prediction can be used for adaptive resource management in wireless systems [10], [13]. Resources in mobile wireless networks are scarcer than in wire-line networks. Wireless links, in general, provide much lower bandwidth than the wired counterpart. This disparity is expected to hold in the near future even though rapid progress is being made for high-speed wireless transmission, due to the physical limitation of the wireless media. Wireless channels are inherently less reliable compared to wired links, mainly because of signal interference, multi-path fading, and shadowing effects. Even with channel coding, diversity combining, and power control techniques, their unreliability is much higher than that of wired links.

The unawareness of the mobile user's future location has contributed in increasing a number of dropped calls during handover. This dropping of calls affects QoS and the overall system efficiency. Much research effort has been put in trying to propose schemes that try to predict the next connection point of a mobile user. Proposals have been made to reserve network resources in the six neighbouring cells around the cell in use by a mobile user [6], [14]. Supporters of these proposals claim the idea works so well in a sense that resources are readily available for a mobile user seamlessly. However, over-reservation of network resources is introduced because a mobile user would not visit all six neighbouring cells at the same time. Other research works propose to reserve network resources based on a mobile user's history [15], [16], [17], [18], [19], [20], [21], [22]. The assumption here is that a mobile user exhibits a regular pattern. Though this idea may work well for users with regular patterns, a high degree of inaccuracy is introduced with random movement patterns.

1.2 Statement of the Problem

Mobile users demand the same reliable service as that present in today's wired networks. *For example, mobile users are likely to* become annoyed if they encounter frequent call droppings or data loss whilst they are communicating, or if they experience any service delay that affects their productivity. Unlike in wired networks, which enjoy the privilege of low error rates and stationary users, there are numerous characteristics of mobile wireless networks, which make it very difficult to provide QoS [2]: We present here three of them.

1. Mobility
2. Limited resources, and
3. Unreliable radio links

Due to these distinct characteristics of mobile wireless networks, it is necessary to develop some mechanisms tailored to support QoS for mobile users. If these networks have prior knowledge of the next connection point of every mobile node in its vicinity, they could have a perfect mechanism for resource reservation in place so that QoS can be guaranteed during the mobile node's connection lifetime. However, such an ideal scenario is highly unlikely to occur in real life. Currently, Mobile IP is based on the best effort delivery model and has no consideration of QoS. Furthermore, the Resource Reservation Protocol (RSVP) model, which is an efficient resource reservation in the wired Internet, becomes invalid under host mobility. The prediction scheme presented is used to predict a mobile node's next connection point in a wireless mobile network.

1.3 Research Motivation

Attempts have been made to develop mobility prediction schemes, which have, either assumed a mobility pattern [23] or road topology-based patterns [24]. The scheme presented here is based on the premise that a future location of a mobile user does not depend on where he/she is currently. This therefore means that a user's behaviour does not necessarily depend on his/her movement history [25]. We set to develop a mobility prediction scheme that utilises real-time MH information. Developing a scheme of this nature is necessary for provisioning of diverse QoS to mobile users. Furthermore, prediction accuracy can be achieved with a considerable decrease in the total tracking area of the network. The scheme presented in this dissertation is efficient for both regular and random users.

1.4 Research Objectives

This research develops a mobility prediction algorithm that will be used to predict the next cell a MH will move into as it moves from one cell to the next during the life time of a connection in a wireless mobile Internet. This next-cell prediction scheme uses a sectored-cell approach. This algorithm is efficient for both regular and random movements. The objectives of this work can be summarised as follows:

(i) To develop a next-cell prediction model, and

(ii) To evaluate the developed model under certain mobility parameters.

1.5 Organization of the Dissertation

The remainder of this dissertation is organised as follows:

Chapter 2 provides the background of the dissertation. Mobility is introduced in section 2.1. An overview of QoS in the mobile wireless Internet is also presented. The impact of mobility on QoS is also described. Mobile IP and MRSVP are introduced. The concept of mobility prediction is also introduced in this chapter. Few schemes that are proposed that attempt to solve the problem of mobility prediction are analysed and presented. Finally, a description of the process of cell sectorisation using directional antennas is given.

Chapter 3 presents a summary of the related work for neighbourhood-based prediction schemes, a history-based prediction scheme, and a background of cell sectorisation.

Chapter 4 presents the sectored-cell approach. A detailed explanation of the partitioning of a cell into sectors is presented. Then, the process of partitioning a radio cell into a critical and a non-critical region is described. Finally, the next-cell prediction algorithm is presented. This chapter deals with the evaluation of the proposed scheme. The simulation model together with all the accompanying parameters are presented. Then, the procedure used in designing and implementing the simulator is outlined. Lastly, the description of the classes used in the simulation process is given.

Chapter 5 presents all the results obtained from our simulation process. These results are then analysed in this chapter. We then highlight some limitations.

Chapter 6 concludes the dissertation. Limitations of the study are also presented. Finally, possible directions for future work are suggested.

CHAPTER TWO

BACKGROUND

2.1 Introduction

It will be desirable in the future mobile wireless Internet to provide resource allocation for the various classes of applications that require QoS guarantees. A severe drop in service quality when a call handoff is experienced while a mobile user moves from one cell to the other may not be acceptable for a number of these applications in general, and real time applications in particular. It is therefore necessary to maintain the required QoS of these applications in the presence of user mobility with the use of resource reservation. QoS provisioning in the mobile wireless Internet has been attempted at various levels in the protocol hierarchy [26]. Mobility of hosts has a significant impact on the QoS parameters of a real-time application. It also introduces new QoS parameters at the connection and system levels. This chapter discusses QoS in the context of the mobile wireless Internet. It also shows how this QoS is affected by the mobile nature of the system. Some mobility prediction schemes are analysed and presented in this chapter.

2.2 QoS Concepts

Majoor R. [26] defines QoS as an idea that transmission rates, error rates, delay, jitter and other network characteristics can be measured, improved, and to some extent guaranteed in advance. QoS defines non-functional characteristics of a system, affecting the perceived quality of the results. In multimedia this might include picture quality, or speed of response, as opposed to the fact that a picture was produced or a response to stimuli occurred. There are two classes of QoS parameters, viz: technology-based parameters and user-based parameters. Technology-based parameters include timelines, bandwidth and reliability, while user-based parameters include perceived QoS, cost and security. This work focuses on the former parameter, that is technology-based parameters. Basic mathematical models for concatenating throughput, delay, jitter, and packet loss rate are specified in [27].

2.3 Impact of Mobility on QoS

Support for user mobility is one unique feature of mobile wireless networks. This is the very same feature that poses a number of challenges in the network. Users tend to move around during a communication session causing hand-offs between adjacent cells. Whenever an active mobile host (MH) moves from one cell to another, the call needs to be handed off to the mobility agent or Foreign Agent (FA) in charge of that new cell, and network resources must be reallocated. Service degradation or even forced termination may occur when there are insufficient resources to accommodate these efforts.

For a better understanding of issues pertaining to integrating a QoS management framework with a mobility framework, attempt was made to focus on QoS managed service provisioning in a wireless Internet using Mobile IP [28], the current proposal for supporting mobility in the Internet environment. Since Mobile IP was designed without the consideration of QoS, the framework has gone through various extensions to accommodate QoS requirements. The major distinction between mobile wireless and stationary networks is that the former has to be able to adapt to the changes in QoS resulting from mobility, rather than trying to provide hard guarantees of QoS [29].

One major problem with respect to mobility is variation in QoS due to handover processes as the MH moves from one cell to the next during an open connection. This period of handover may result in a short loss of communication, which may not be noticeable for voice interaction but can result in loss of data for other applications. Another critical problem is that of choosing an appropriate base station to which an MH will handover, and this is the problem this research is trying to solve. Also, wireless networks have blind spots behind buildings or hills, under bridges, where the signal may be very weak resulting in temporary quality reduction or connection loss when an MH is in a moving car or train. Inconsistencies in wireless link quality may also be as a result of atmospheric conditions such as rain or lightning. These effects require a more sophisticated QoS management than in stationary wired networks. It can therefore be deduced that the difference between the wired and wireless networks is in the variation in QoS.

The effects of mobility on QoS require that all schemes to be employed be capable of managing all network-based parameters (that is timelines, bandwidth and reliability) mentioned above. Equally important, all these should be effected at the lowest possible costs.

2.4 Mobile IP

Mobile IP is the technology that allows the "mobile node" (MN) to change its point of attachment to the Internet while communicating with the "correspondent node" (CN) using IP. The goal of Mobile IP is to provide mobility support for hosts connected to the Internet without changing their IP addresses. The mobile host's current care-of-address (e.g., the IP address of a base station, called a foreign agent (FA)), to which it is currently connected) is registered with its home agent, which is attached to its home network. The HA then intercepts and forwards its to care-of address by encapsulation [30]. The FA then decapsulates the packet and delivers it to the MH. A detailed description can be found in [31].

Mobile IP suffers from triangle routing: while the MH can send out packets through the FA and along an optimal path, incoming packets have to travel through HA. If the current location of the MH is close to the sender's but the HA is far away, packets have to take a long detour (see figure 2.1). Mobile IP route optimization [32] provides a mechanism to alleviate this problem. Any host, which is willing to participate, maintains a binding cache. When the HA intercepts a packet for an MH that is away, it may send a binding update message to the source of the MH's current care-of-address. The source then updates its binding cache, and tunnels any ensuing packets for the MH directly to its care-of-address.

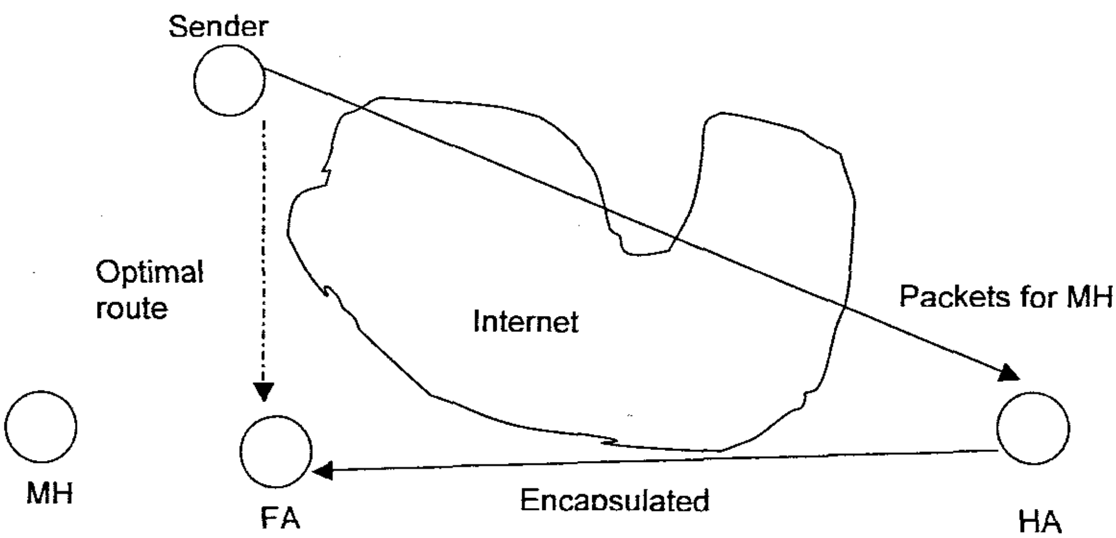

Figure 2.1: Triangle Routing Problem

FAs can make use of binding updates to reduce packet loss during handoff. As an extension of the registration process, the MH may ask the new FA to send a Previous Foreign Agent Notification message, which includes a binding update, to the previous FA. The previous FA then updates its binding cache and re-tunnels any packets for the MH to its new care-of-address. This process is called smooth handoff. These forwarding mechanisms, however, open a potential security hole for the Internet hosts. Without Mobile IP, a malicious host may have to physically break into a network to grab packets for another host. To eliminate these security exposures, Mobile IP requires that registration requests be authenticated. An identification field in registration messages provides replay protection [31]. Authentication procedures are also required in route optimisation for protection from similar attacks. A source host that accepts binding update messages must authenticate them. During a smooth handoff, the previous FA must authenticate the binding update message using a registration key established with the MH.

2.5 Mobile Resource Reservation Protocol (MRSVP)

The mobile resource reservation protocol (MRSVP), proposed by Talukdar *et. Al* [16], is an enhanced version of the RSVP used in wire line IP network. During the life time of the connection, the protocol can make use of the concept of proxy agents to make advance resource reservation along the data flow path to and from the locations it may visit. First it needs to specify the set of locations, called MSPEC, from which a mobile host wishes to make advance reservation. The proxy agent at the current location is called a local proxy agent and the ones at the other locations in its MSPEC are called remote proxy agents. The remote proxy agents will make passive reservation on behalf of the mobile host. The local proxy agent will act as a normal router. The issues arising out of this are those of determining who will be the proxy agent for the particular mobile host, (the process is called proxy discovery) and how the MSPEC file is generated i.e. the problem of determining in advance, the set of locations a mobile is likely to visit. The MRSVP assumes that the mobile has somehow acquired its mobility specification file. This file can be modified dynamically.

MRSVP specifies two types of reservations: active and passive. The active reservation is initiated from the current location and the passive reservation is from its future location. To improve the utilisation, the passively reserved bandwidth would be used by other flows that are occurring at the location at that time, and which require a relatively weaker QoS guarantees. But this flow may get affected once this passive reservation becomes active (i.e. when a mobile moves into that location) since it will have to surrender the bandwidth. To avoid this, the resources of the passive reservation are multiplexed among different classes of users.

The mobile host selects its proxy agent by using the Proxy Discovery Protocol. A proxy agent is a special MRSVP capable router that delivers all incoming and outgoing packets to the mobile. It also makes passive reservations on behalf of the mobile in the remote location. Besides, it also informs the mobile of the results of the reservation attempt for the passive reservation it tries to make.

2.5.1 MRSVP Protocol description

A brief overview of the MRSVP protocol is given below but the detailed description is available in [16]. The MRSVP protocol is assumed to run on top of the Mobile IP. This protocol also assumes that the mobile is aware of its mobility configuration file beforehand via some external means that this protocol does not deal with. The protocol functions as follows:

i. The sending node sends a PATH message to the mobile node on its unicast address or the multicast address that the mobile is a part of;

ii. Mobile host responds by sending its mobility specification file to all the potential foreign agents that are listed in this file.

   a. In case of unicast flow reservations, the mobile host also sends this file to its home agent;

   b. In case the reservation is for multicast flow, the mobile host will ask the agents listed in the mobility specifications, to join the multicast group; and

   c. Then, it will send RESV message to the sender to make an active reservation. All subsequent receipts of the PATH messages by the mobile node from hereon, will only prompt it to send the RESV message to the sender.

iii. On receiving the mobility specification file, the home agent will tunnel all PATH messages meant for the mobile host to all the mobility agents in the specification file.

iv. On receiving the PATH message from the mobile host not in their network, the mobility agents will send a RESV message to the sender to make passive reservation.

v. On detecting a change in the flow rate, the mobile node may send new flow specifications to all the mobility agents in its mobility specification file.

The advantage of MRSVP is that the required resources for the mobile host in the new region can be retrieved rapidly because the resources will have been pre-reserved in the original passive reservation path. Secondly, seamless handoff for QoS guarantees can be retained using the MRSVP protocol. However, MRSVP wastes considerable bandwidth in making advance resource reservations. This excessive resource waste may degrade the system performance. Another drawback is that mobile node originating for real-time flows has to wait until all the necessary resources in a possibly large set of attachments become available. Therefore, eventually the blocking rate of the flows originating for real-time applications may become very high. It is also difficult to accurately determine the possibly large set of attachment points of a mobile host.

CHAPTER THREE

RELATED WORK

3.1 Introduction

This chapter gives a background of mobility prediction. It also presents mobility prediction schemes that are proposed in the literature. These schemes are classified into two categories according to the way their prediction is done. The categories are neighbourhood-based reservation schemes and history-based reservation schemes. Finally, the way handover should be done in the sectored-cell structure is discussed.

3.2 Mobility Prediction

It is non-trivial to determine where to reserve network resources due to difficulties with regard to user movement prediction. Despite the challenges, many mobility prediction schemes have been proposed in the literature as an attempt to safeguard QoS agreements of mobile services. According to Chan and Seneviratne [23], these algorithms can be classified into two categories, namely the neighbourhood-based prediction and the history-based prediction as discussed in the next subsection.

3.2.1 Neighbourhood-based Prediction

Schemes falling under this category reserve network resources between an active cell and a set of base stations surrounding the mobile node, Figure 3.1. The number of base stations involved in the reservation process depends on how long (service time) the network is willing to support a mobile service. For example, Olivierra *et al* [6] proposed an adaptive bandwidth (resource) reservation scheme whereby a new call or a handoff call is accepted only if it can successfully reserve resources in all neighbouring cells. This is in addition to the resources already available in the local/active cell.

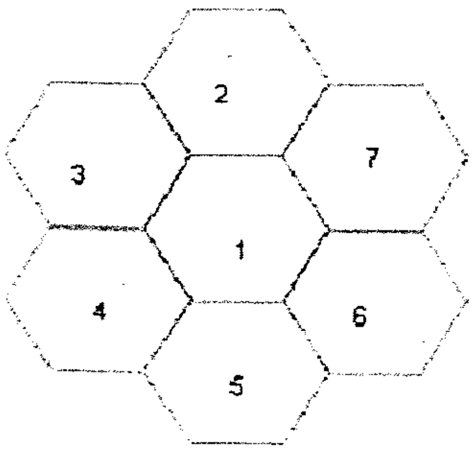

Figure 3.1: Seven-cell Cluster

This over-reservation of network resources in all neighbouring cells is inevitable, since there is no consideration for the individual trends of the users in the network, e.g. *position, speed,* and *direction.* Reservations are made even when the MH is not mobile (stationary). In addition, each MH has to reserve resources in its neighbouring cells the moment it enters a cell. This inflexibility in reservation location and timing could lead to under-utilisation of network resources. Also, in Advanced Reservation Signalling [14], resources are reserved only between the current location and neighbouring cells. Thus the network guarantees continuity of services after the next handover, but its further commitments are subject to successful reservations at the new neighbouring cells.

3.2.2 History-based Reservation

Schemes falling under this category use the user mobility history to predict the future movement of a mobile host. Depending on the service commitment to mobile users, these proposals reserve resources at various levels in advance in the predicted path. For instance, to obtain mobility independent service guarantees, the MRSVP and other similar proposals [16], [17], [18] attempt to make resource reservations at each cell a mobile host may visit during the lifetime of a connection.

In [3], the authors propose the concept of a shadow cluster: a set of base stations to which an MH is likely to attach in the near future. Based on the probabilities of a visit in the past and the current trajectory of a mobile host, network resources are reserved near its present location and along its direction of travel. The scheme partitions time into equal intervals, and estimates the probability of each MH being in any wireless cell within the shadow cluster for future time intervals. During each time interval, the base stations exchange information about the predicted bandwidth demands for future time intervals so as to determine the feasibility of admitting new call requests. The scheme assumes precise knowledge about individual MHs dynamics and call holding patterns in the form of probability density functions (pdfs), which an MH needs to submit to the base station at the time of making a new call request. Such user-specific information may not be available in all MHs; it is not clear how the scheme would handle such cases. Another drawback is that resource reservation can be excessively aggressive causing an increased blocking of new calls.

Choi and Shin [19] proposed another Predictive and Adaptive Bandwidth Reservation Scheme. Bandwidth resources are reserved in neighbouring cells according to the estimated probability that an MH would handoff into these cells within an estimation time window, based upon the MHs previous cell, and its extant sojourn time (time already spent in the current cell). It requires the use of a knowledge base containing the time spent by previous MHs in the cell, the previous cells that they came from, and their corresponding target handoff cells. However, it may be insufficient to predict the mobility of an MH based on its previous cell information and its extant sojourn time. Moreover, calls that are newly generated in the cell do not have previous cell information. This may further reduce the schemes prediction accuracy.

A less ambiguous resource reservation scheme can be found in the Profile Based Next-Cell Prediction [20], where network resources are reserved only at the most likely visited cell, and further QoS commitments depends on the reservation process after the next handover. It is noticeable that the more a scheme tries to predict the movement, the more likely a network will support a lifetime of a session. However, this is achieved at the expense of the overall network utilisation because of poor prediction accuracy. Also, movements are restricted to indoor locations such as an office, corridor or common room.

The Regular Path Recognition Method [21] attempts to exploit regularity in human behaviour in terms of periodic delay activities such as travelling to work, travelling to school, etc., which results in probabilities that can be assigned to use paths. The more the use of recorded cell patterns is made, the more likely the path of the user is detected. In a way this mobility prediction method is an extension of the mobility prediction model making use of the segment criterion and suffers from those same drawbacks. The accuracy of the path detection depends on the amount of user profile data available. This is also assumed that all user movements can be contained as a regular path.

The Mobile Motion Prediction scheme [22] makes use of the user's movement history to make path predictions. Movements are considered to be a combination of random and regular movements and are merged using a Markov chain model made up of movement patterns. The scheme depends on the availability of the user's movement history. As a result the performance of the Mobile Motion Prediction scheme is accurate with regular movement patterns but decreases linearly as random movement increases. The Hierarchical Position Prediction scheme [8] makes use of the user's history of movements together with the instantaneous RSSI measurements of surrounding cells. While agreeing that the next cell mobility of the mobile user is governed by the movement pattern of the user within the current cell area, but we argue that the tracking of the mobile user need not be performed in the complete cell area. This is due to the fact that, users near the centre of each cell have comparatively low chances of going to the neighbouring cell. This mobility prediction scheme remains reasonably accurate despite the influence of the random movements.

3.3 Cell sectorisation in wireless networks

Sectorisation is a technique of using directional antennas for transmission and reception [33]. The 120-degree sectoring is usually employed and the capacity increases three fold, (figure 3.2). Recently, as cell sectorisation technologies become more promising, some newly proposed schemes take advantage of directional or smart antennas to partition cells for capacity improvements. This could potentially give rise to better prediction accuracy and greater adaptability to time-varying conditions than previous methods. Due to antenna directivities, power voltage is reduced significantly in other directions; therefore interference from the side and back lobe is very minimal [34]. The impact of cell sectorisation on the performance of wireless networks can be seen on a number of ways: Firstly, it is due to the narrower beam of sector antennas resulting from less interference received.

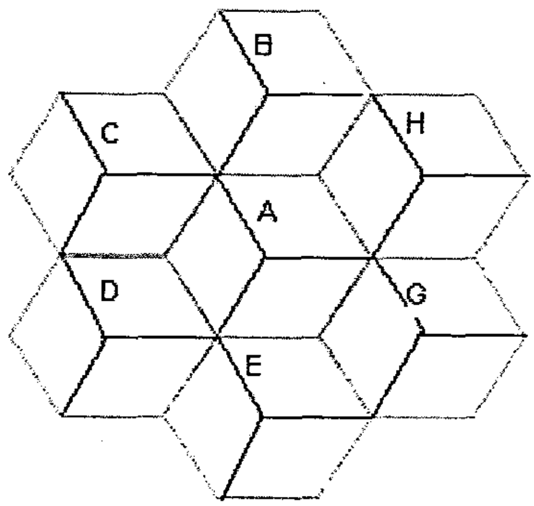

Figure 3.2: 120-degree Tri-sectored Cells

Furthermore, the higher gain of the sector antennas allows for a reduction in the base station transmit power leaving the power amplifier and therefore allowing more users to be served with a fixed maximum total power.

3.4 Mobility Prediction-based Handover

This section discusses handover as a basic mobile network capability for dynamic support of mobile terminal migration [35] and covers its basic requirements. An efficient bandwidth reservation scheme for reserving network resources in the predicted or target cell(s) is necessary to avoid unacceptable forced termination of calls and waste of limited resources in the wireless mobile communication. Directional antennas have this in-built functionality of handling handover between sectors belonging to a single cell; hence this work focuses on cell-to-cell handover in the presence of sectorisation.

Hansen *et al* [36] identified two different ways of implementing a handover: backward and forward handover. In the first technique (forward handover), an MH reports the need for handover and the target access point it can hear better than the current one, and thus is able to decide when to hand over to another access point. The major drawback of this criterion is seen when a wireless network is touched around its blind spot mentioned above. This results in the execution of unnecessary handoffs. In the second technique (backward handover), new access points are not pre-selected. In this kind of handover, the MH will first release the old radio link and handover will occur only after it has chosen the new access point. In this case the handover request towards the network is sent through the new access point after the MH has established a radio link.

3.4.1 Basic Requirements of Handover

A good handover management scheme should be based on the following design factors for it to be acceptable in the wireless industry [35]:

i. *QoS* – should be maintained or re-negotiated;
ii. *Latency* – the time required to effect the handoff should be appropriate for the rate of mobility of the MH, as well as the nature of data transferred;
iii. *Scalability* – the handoff procedure should support handoffs within the same cell, between different base stations in the same or different networks and
iv. *Minimal drop-off* and *faster recovery.*

Assuming that the prediction scheme succeeds in making accurate predictions, the outstanding question will be: how is the process of handing calls off to the predicted cell carried out? Another issue is that of ensuring that if resources are not available at the predicted cell(s), calls are not forced to terminate but to still be admitted in the predicted cell. This can be made possible by employing a handover strategy that will allow the borrowing of channels from new calls channels for the purpose of handover calls, or in the second nearest cells given that signal strengths received from that cell are enough, to serve the incoming MH. When an MH attempts to handoff from one cell to another, it may encounter forced termination due to a resource shortage at the target cell. The probability of this occurring is often referred to as the forced termination probability (P_{FT}) of a handoff.

From a mobile user's point of view, forced termination of an ongoing call is more annoying than the blocking of a new call request. Therefore, handoff-requests are generally prioritised over new call requests when they compete for radio resources. In the classic handoff prioritisation problem, each BS prioritises handoff requests by setting aside some wireless resources that could only be utilised by incoming active calls. Since any such resource reservation would inevitably increase the blocking probability of new calls (P_{CB}) and reduce the overall system resource utilisation, it is therefore extremely important that these reservations are made as sparingly as possible, while meeting the desired probability. In this way, wireless service providers would be able to provide high quality services without compromising their revenues unnecessarily.

Early work on handover prioritisation proposed static reservation of resources at each BS as a solution [37], in which a fixed portion of the radio capacity is permanently reserved for handoffs. However, such a static approach is unable to handle variable traffic load and mobility; it might under-utilise precious radio resources when handoffs are less frequent, and could experience unacceptably large number of forced terminations when the mobility is high. In order to meet the desired P_{FT} without over-reserving precious radio resources, the amount of reservation at each BS should be dynamically adjusted according to the requirements of anticipated handoffs.

CHAPTER FOUR

MODEL DEVELOPMENT

4.1 Introduction

While there has been previous work in the literature that attempted to perform next cell prediction based on user mobility history [16], [17], [18] and based on neighbouring cells [6], [14] they all have underlying assumptions that yield unrealistic results and also lead to over-reservation of resources. This chapter presents a mobility prediction scheme that utilises real-time mobile host parameters such as current position, direction, and speed in a particular cell-sector. The scheme requires the serving BS to receive regular updates about each active MH's position every time *t*, say 1 sec. This will consume a small amount of uplink wireless bandwidth (several bytes per update for each MH), which might be negligible for future broadband services. An important point to emphasise here is that sectors and the positioning information of the MHs are only used to predict the next cell an MH will move to, so as to make dynamic adjustment of resource reservations; actual handoff-requests will still be initiated based on received signal strength measurements. The next section discusses the partitioning of a radio cell.

4.2 Sectored Cell Structure

This is a simple mobility prediction scheme. The key idea is to predict an MH's next connection point using critical region information and its most current positions in a particular sector. The objective is to predict the next cell a mobile host will move into as it moves from one cell to the next with an open connection. In order to perform accurate predictions, each cell is partitioned into sectors and then the critical region is estimated as discussed in the next section.

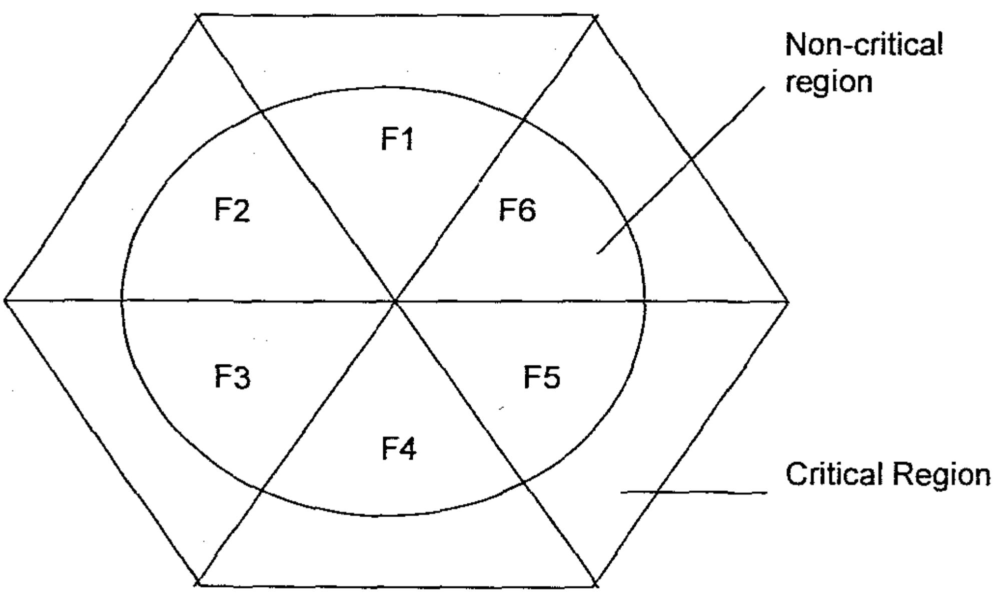

Figure 4.1: Sectored-Cell Structure

While most conventional cellular systems employ three 100-degrees to 120-degrees directional antennae at each base station [34], in this work each radio cell was partitioned into six sectors using directional antennas as shown in figure 4.1. A hexagonal cell structure was assumed. Six directional antennas were employed to partition a region around each BS into six sectors. The partitioning was done such that each sector is adjacent to each neighbouring cell.

4.3 The Critical Region

After partitioning a cell into sectors, the next task was to determine the critical region (CR) of each cell. The CR was obtained by demarcating a region with its internal boundary being the average distance from the antenna site at which handoff-requests usually take place. These are all the points between where a number of handoff-requests were made and where real handoffs occurred. This concept is shown in figure 4.2.

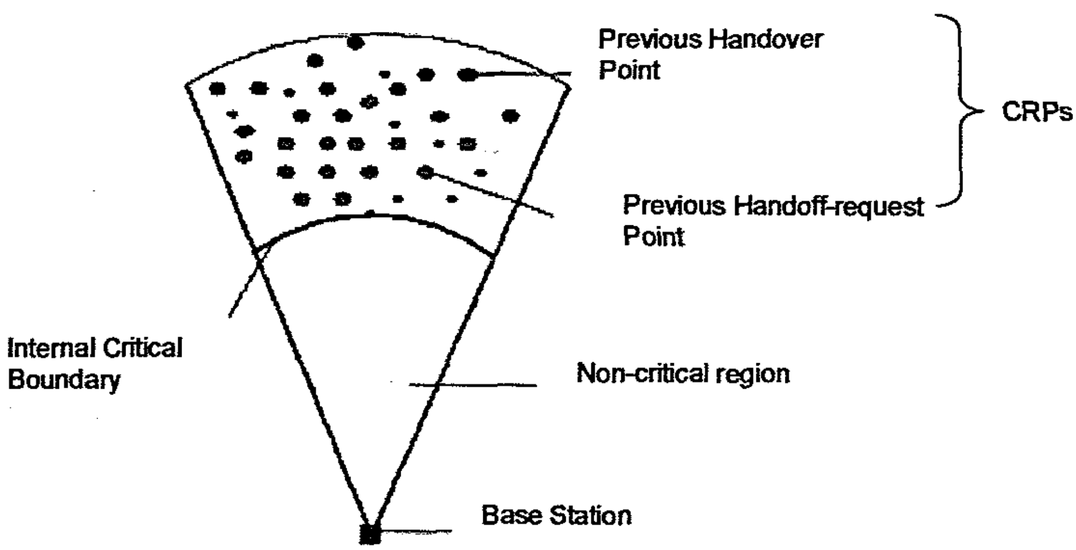

Figure 4.2: Critical Region Representation Points (CRPs).

These points are referred to as *critical region representation points* (CRPs). In order to compute the CRPs, the BS had to collect data about the locations where previous handoff-requests were made, as well as the corresponding target cell for each such request. It bears noting that a MH's target cell would already have been chosen by the time it makes a handoff-request. All CRPs fall into two imaginary borders referred to as internal and external critical borders respectively. The collections of handoff-request positions are divided into *groups* according to the sector in which they are found. The most likely target cell is also associated with each CRP, if it happens to be the most commonly chosen target cell among all the handoff-requests that occurred within the sector. The CRPs and their corresponding most likely target cells are stored in a table located at each BS. This table is updated, say, once a day, based on handoff-request positions collected over the previous day. In this way, they are robust against changes in terrain characteristics or man-made structures that may affect radio propagation, as well as modifications in handoff parameters (e.g., hysteresis margin, handoff thresholds [38]) by the network operator.

The number of entries in the table was exactly equal to the number of CRPs, *N*. The parameter *N* was allowed to differ for each BS; it was chosen experimentally to ensure that they are indeed good representation of the CR. The choice of *N* should depend on the size of actual BS coverage area, as well as, the topography; a larger BS coverage area or the presence of more irregular topographical features will require a larger *N*, so as to increase the resolution of the represented critical region.

4.4 Mobile Host Parameters

The next-cell prediction scheme utilised sectored-cells together with mobile parameters such as the current position, direction and speed of the MH. The use of real-time MH parameters for mobility prediction gives rise to better accuracy and greater adaptability to time-varying conditions than previous methods.

4.4.1 MH Current Position

The current position of the MH plays a very important role in the proposed scheme. Dru and Saada [39] say that it is possible to refine positioning using the measurements taken by the base station, which measures the time between the transmission of a frame (from the base station to mobile) and the reception of the corresponding frame (from mobile to base station). Using this measurement, the base station can work out the distance to the mobile, with considerable theoretical accuracy. This makes it possible to restrict the area of inaccuracy. Although the information is of little use for a cell served by an omni-directional antenna, it can offer improved accuracy in the case of sectored-cells served by several antennas. As mentioned earlier, the BS obtained position information at regular time intervals *t* (typically 1 sec). It also stored the MHs previous positions over a specific number of intervals for the purpose of estimating speed and direction of the mobile.

4.4.2 MH Current Speed

The MH's current speed plays a very important role in the proposed prediction scheme. The assumption here was that, the probability that a fast moving mobile node would handoff to the predicted cell is very high, due to the fact that the likelihood of it changing a direction as fast as it could, is very small. For a slow moving user, the scheme always has enough time to predict and re-predict in cases where the user changed direction. We used the simple theory of vectors to estimate the MH's current direction. It is proven that, if at least two points of any moving object can be determined, then the time taken by an MH to move from one point to the next can be determined. The speed (v) can be calculated as $v = s/t$.

4.4.3 MH Current Direction

The MH's direction of travel was estimated using vectors drawn over these X points. Note that $X > 2$ is required to mitigate the effects of random positioning errors. However, X should not be too large, else the algorithm could become insensitive to real changes in direction.

4.5 Prediction Algorithm

Instead of delegating the prediction responsibility to individual MHs, this task was assigned to base stations. This reduces the computational power requirement at the MHs, which is more attractive since battery power limitation and component cost are major concerns for mobile device manufacturers. Also, the BS is able to handle this additional computational requirement without any difficulty.

Predictions were only performed for MHs that were within the critical borders (inside the CR). These were the MHs that had the greatest potential of making handoff-requests given the fact that the CRs are represented by previously occurred handoffs and handoff-requests. As mentioned earlier, we assumed that the BS is able to obtain the MH's current position information every time interval *t*. The BS kept a record of the most current positions of each MH. That was used to obtain the MH's average speed, as well as to determine the MH's position. In the next-cell prediction algorithm shown in Table 4.2, the description of how mobile positioning is performed is not given. Instead, it was assumed that the MH's current *X positions* could be obtained using other positioning techniques, e.g. GPS [41].

During the prediction phase, the other two parameters (direction and speed of the MH) had to be specified. These parameters together with the CRP associated with the current position of the MH were used to calculate the probability of an MH handing-over to a predicted cell *C* if the MH were to handoff within that sector. They are dependent on the

MH's current position within the cell-sector; therefore they have to be recomputed during each prediction.

For every active MH in the cell, the BS performed a series of simple checks to estimate the current location of the MH, whether the MH was in the CR or not, based on the represented critical region information. If it found that the MH was not in the critical region, the check method returned false and no prediction was made. When this happened, the system did not reserve any resources for this particular MH.

In cases where it was found that an MH was in the CR, its direction was computed to check if the MH was departing (that is migrating away) from the cell. If the result was true, which implies that the MH was departing, the CRP, which was closer to the MH's current location, was determined. The corresponding target cell was therefore predicted to be the cell *C* associated with the closest CRP (within a particular sector), while the corresponding probability that cell *C* was indeed the target cell was computed. The probability was computed as a function of the CRP associated with a specific position, speed and a direction of the MH. This corresponding probability was necessary for two reasons: firstly, we computed this probability so as to know how much resource should be reserved in the corresponding target cell. Secondly, there was a need to safeguard against the possibility of handoff failure due to insufficient resources in the predicted cell. With the probabilistic information, another cell, which was implicated by the probability, was used to accommodate the departing MH. Table 4.1 gives all the notations used in the Psuedo-code

Table 4.1Notations used in the Pseudo-code

Definition:
$S_{(i,j)}$ is sector identifier
Where: $S_{(i,j)} \in \{1,..,6\}$, sectors in each cell
I is the current cell
J is the target cell
P_C (x) = the probability that the mobile Node will move to cell C
CR = critical region of the current cell
NCR = non-critical region
CP = current position of the MH
CRP = critical region representation point
P_T = threshold probability

Table 4.2: NCPS Pseudo-code

```
The pseudo-code is defined as follows:

IF MH is in critical region THEN

    If MH is departing   THEN
          Associate MH current pos with closest CRP
          Compute P (x)

          Prediction:
           IF Current sector = S(i,j) AND P(x) > p     THEN
              Target cell = Cell-j
              Reserve resource in Cell-j
            END IF
    END IF;
ELSE Target cell = current cell
        Monitor mobile
END ELSE
```

P(x) was calculated by taking into account the CRP associated with the current position of the MH, its direction as well as the speed at which the MH was travelling. Monitoring the time interval depended on the size of the cell.

4.6 Prediction Efficiency

For the purpose of the reservation algorithm, the efficiency of the prediction *(PE)* was also considered. The efficiency of any prediction scheme is calculated from the number of cells that particular scheme is able to predict out of six (if we assume that six cells surround the current cell). This means that for a scheme that is reserving resources at all neighbouring cells, the efficiency ratio of that scheme is 0%. This is simply because there was no prediction made.

At the other end, if a prediction scheme is able to accurately predict one cell out of six (1/6), it means that this scheme is perfect and the efficiency ratio is 100%. From the three possibilities presented below, we can be able to make some logical deductions with respect to efficiency ratios of our proposed scheme.

The formula for calculating the efficiency ratio is:

$$ER = \frac{N_T - N_P}{N_T - 1} \qquad (4.1)$$

Where:

ER is the Efficiency Ratio

N_T: total number of neighbouring cells

N_P: total number of predicted cells

With a prediction that uses a maximum of two neighbouring cells, therefore we have:

$$ER = \frac{N_T - N_P}{N_T - 1} = \frac{6-2}{6-1} = \frac{4}{5} = 80\% \qquad (4.2)$$

This means that the proposed scheme, at worst, would be 80% efficient.

From figure 4.3 it can be noted that a mobile node in sector f1 of cell A, while still in the non-critical region of cell A, will be receiving signal with an ID {A, f1}. Once it enters the critical region of cell A in sector f1, depending on its position, the mobile node may be receiving signals from:

Both {C,f4} and {B,f3} ....possibility #1
Both {C,f4} and {D,f5}, or..................possibility #2
{C,f4} only..possibility #3

If a mobile node is receiving the signal with the identity {C, f4} and {B, f3} as depicted in the first scenario, the system will reserve resources in cells B and C. If a mobile node is receiving the signal with the identity {C, f4} and {D, f5} as depicted in the second scenario, the system will reserve resources in cells C and D.

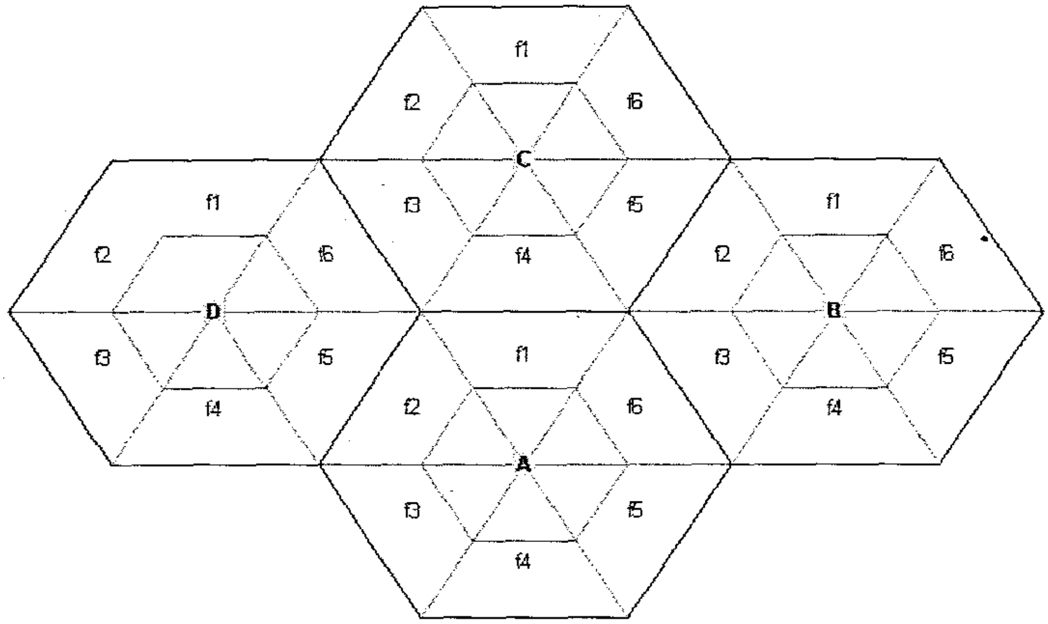

Figure 4.3: Sectors

If a mobile node is receiving the signal with the identity {C, f4} only, as depicted in the third scenario, the system will reserve resources in cell C only. The base station would already have knowledge of all the MHs present in the cell. We assume that neighbouring id {A, f6} and {A, f2} will be ignored to keep the process simple. This assumption is reasonable since these sectors are also controlled by base station A.

In a cellular environment, each cell cluster consists of seven cells; one in the middle and six neighbouring cells. It is also assumed that the resident cell of the mobile node is always at the centre of the cluster. If a mobile node moves from cell A (which is the centre) to cell B, cell B becomes the centre.

An important point worth mentioning is that if one is developing a scheme of this nature, one needs to make a provision for mobile terminals to alternate between nearest base stations from which it is receiving signals. At the same time we do not want to make reservation in all six neighbouring cells. So, we strongly believe that the MH could be allowed to alternate between two base stations, if the combined probability that it might move into them is very high

4.7 Simulation Model

To facilitate the evaluation of the NCPS, a simulation model was carried out. Previous work in the literature either assumes that mobile users exhibit fixed or regular mobility patterns or assumes that mobile users follow random movements. The simulation model used in this dissertation assumes both regular and random movements. This established a realistic platform to evaluate the performance of any mobility-based prediction scheme.

The simulation network consisted of 4 wireless cells as shown in figure 4.4. Each radio cell was partitioned into six sectors. Each sector was then divided into a non-critical region and critical region. In order to eliminate boundary effects that could make it very difficult to comprehend the performance evaluation results, a common approach found in the literature was used. Cells at the boundaries were wrapped. In this way, whenever an MH travels out of the network boundary, it was re-injected into the network again via the appropriate wrap-around cell as though a handoff had occurred from outside the simulation environment [40]. This compensated for any traffic loss at the network boundary. One correspondent host (CH) was placed outside the network so it could continuously send packets to corresponding MHs. This was done solely to avoid any packet loss due to false handovers.

Although the cell layout adopted was a hexagonal cell model (Figure 4.1), the simulation model did not assume that handoffs occur at the hexagonal boundary.

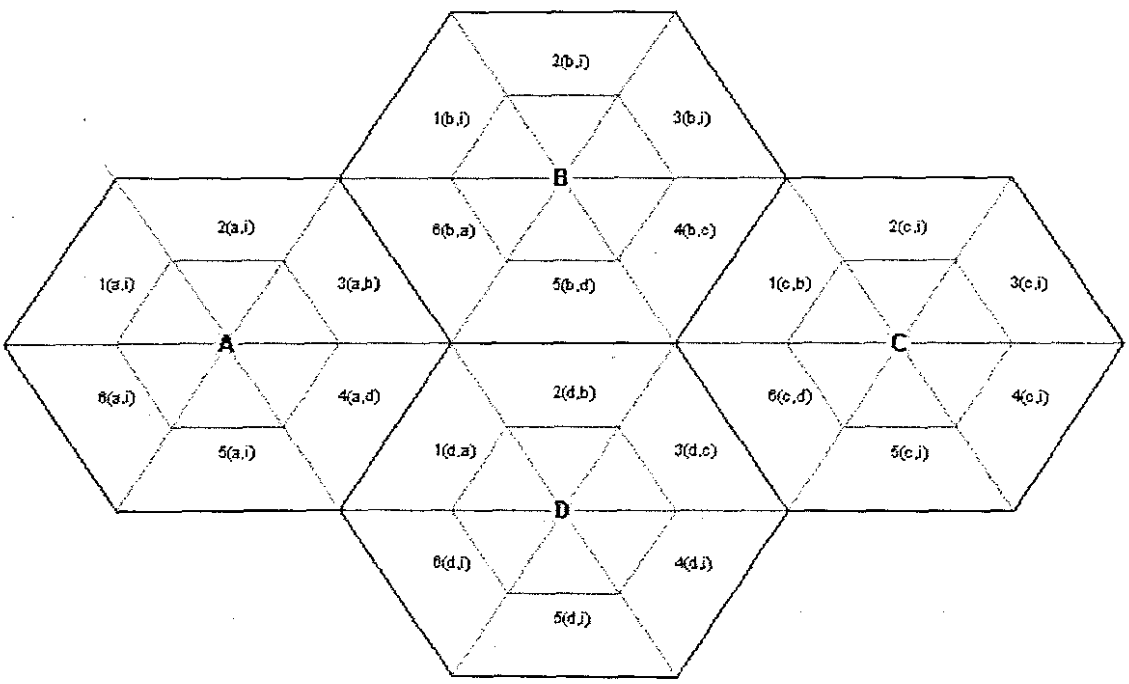

Figure 4.4: Simulated Network Model

Let R be the cell radius (assumed to be 1km in our simulation), defined as the distance from the BS to the vertex of the hexagonal cell model. For the NCPS model to bring out any meaningful information, the output was determined. For any number of predictions, we kept count of the number of accurate predictions made and, the total number of packets lost due to false prediction or due to insufficient amount of available resources in the target cell.

4.8 Simulation Parameters

Two parameters, the MH speed and the size of the critical region, were varied during the simulation. It is to be noted that the MH was allowed to move in any arbitrary direction (between 0 and 360 degrees). The speed of the MH is a random variable. The minimum speed was assumed to be an average speed a slow-moving user can walk, which was between 5km/h and 20km/h. The medium speed was between 20km/h and 80km/h. Highly mobile users were between 80km/h and 140km/h to accommodate fast-moving vehicles. We did not assume any positioning technology (such as GPS) for the MHs, as new breakthroughs in such technologies continue to surface.

During simulation, the boundary between the critical region and the non-critical region was assumed to be at 60% of the cell radius. This parameter was altered so as to study the effect brought about by the size of the critical region with respect to the prediction scheme. The target BS was assumed to be the neighbouring cell adjacent to (facing) the sector containing the MH. It was also assumed that the smallest amount of bandwidth that may be assigned to any connection is 1 Bandwidth Unit (BU), which was taken to be the minimum required bandwidth to support voice connection. Each cell was assumed to have a fixed link capacity C of 100 BUs. Therefore, a single sector had a minimum of 100/6 BUs, assuming that by partitioning a single cell into sectors; a cell capacity is also increased. The bandwidth requirement of each MH was assumed to be symmetric, meaning that they have the same requirement in both uplink and downlink. However, it is straightforward to modify the scheme to handle asymmetric requirements. All MHs were assumed to have the same probability of forced termination requirement, regardless of their connection types. All communications here were assumed to be real time communication with high level of QoS demand.

4.9 Simulation Design and Implementation

All classes that make up our simulation were written in the JAVA programming language using Jbuilder5 environment. We chose JAVA over other programming languages simply because JAVA is an Object Oriented Language. The simulator had a visual component that allowed the experimenter to continuously view the network for the duration of the simulation. This component had been very useful during the implementation of the simulation. The first step of identifying objects and classes began by generating a set of candidate classes and objects using the classical approach and behaviour analysis. The eligible classes and objects generated using the classical approach are summarised in Table 4.3. Abstractions that relate to system function points revealed during behaviour analysis are shown in Table 4.4.

Table 4.3: List of eligible objects discovered during analysis

Tangible Things	**Roles**	**Events**	**External Systems**
Base Station	Mobile User	Predicts cell	Service Providers
Antenna	Mobile Operator	Handoff	Tall Buildings
Mobile Device	Mobile User	Request handoff	Bridges
Cell	Mobile User	Send message	Tunnels
Network	Mobile Operator	Create links	Telecommunica tion Companies

Table 4.4: List of behaviours discovered during analysis

Behaviours
1. The system checks if an MH is in the critical regions 2. The system checks if an MH is departing 3. The system computes the MH direction 4. The current BS predicts the next cell using MH parameters 5. The predicted cell reserves resources for an MH 6. The system checks an MH requirements versus target cell capacity 7. An MH is handed over to the target cell

After perusing the list of possible objects, the base station (BS) was determined to be a primary object. The roles and responsibilities that this abstraction should encompass were considered. The BS is responsible for keeping track of all MHs in its vicinity, the direction they are travelling, the state of operation, and keeping appropriate information about all neighbouring base stations. Parameters such as direction, speed, and state may seem like overkill for a simple prediction scheme. However, in anticipation of possible reuse or expansion of the program, they were included. The rest of the design and implementation is described next as (i) class description and (ii) user interface

4.9.1 Class Description

We used eight classes as shown overleaf (Figure 4.5).

1. **The SectoredCell:** this class has some bandwidth and knows its 6 neighbours. It keeps track of the MHs resident within it, and can answer request to acquire, reserve, or release bandwidth. It keeps the statistics that are compiled at the end of the simulation to measure the various QoS parameters being studied.
2. **The Network Class:** this class contains all Cell objects. hexagonal cells were used to construct the network. The

Network object knows where it is and knows all its neighbours as well.

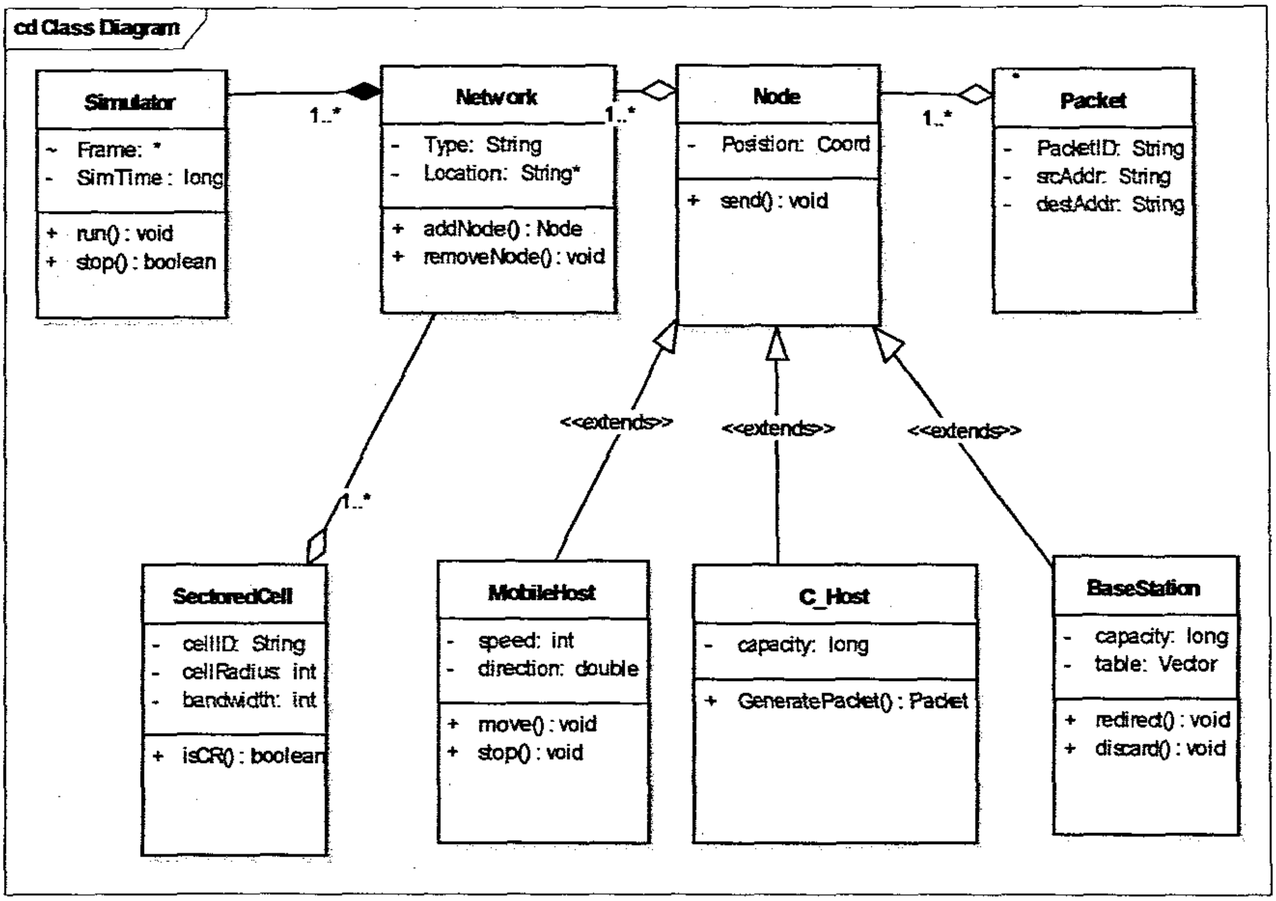

Figure 4.5 NCPS Class Diagram

3. **The Node Class:** an abstract class, which was a fundamental building block of the network. By sub-classing, we made a base station, a mobile host, and a correspondent host class, that were used in the same simulation.
4. **The MobileHost:** this class was given various attributes like speed, direction and bandwidth needs. This class enabled the MH to move randomly within the network. A mobile host was

able to receive packets while it was roaming from one cell to the next.

5. **The C_Host:** This class was included in the simulation so as to allow for the packets to be sent to the mobile host. It also extends the **Node** class.
6. **The BaseStation:** This class also extends the **Node** class. It was responsible for keeping information regarding all mobile hosts within its vicinity. It was also responsible for receiving and relaying (or even discarding) of packets from a correspondent host to a mobile host. In this class we implemented the next cell prediction algorithm.
7. **The Simulator class:** this class has a separate thread of control. It created the objects that participated in the simulation, set them in motion, updated the display, and then disposed of the players after the simulation time expired.
8. **The Packet Class:** this is a simple class that represents a network packet. It only has source and destination addresses as parameters.

The sequence of actions undertaken by various network entities as well as the inter-network signalling required for their completion is depicted in the NCPS Message Sequence Diagram (Figure 4.6). Here, we first assumed that an MH is in a critical region. The current cell checked if the MH was departing. The **DetermineTarget [MH_ID, BS_ID]** signal was used for triggering the prediction algorithm in the current BS. Its parameters denote the identification of the MH for which the algorithm should execute, as well as the identification of the current BS. The current BS invoked the next cell prediction algorithm and notified the current cell, through the **RelocateResource [MH_ID, Direction, TargetBS_ID]** signal. The TargetBS_ID parameter is a list that contains the identifier of the neighbouring cell (to that currently used by the MH), which has a high probability that it is a target cell. The current cell forwarded an MH bandwidth requirement to the target BS using the **MH_Rqmnt [MH_ID, MH_Rqmnt]** signal. Upon receipt of this message, the target BS reserved the required amount of resources.

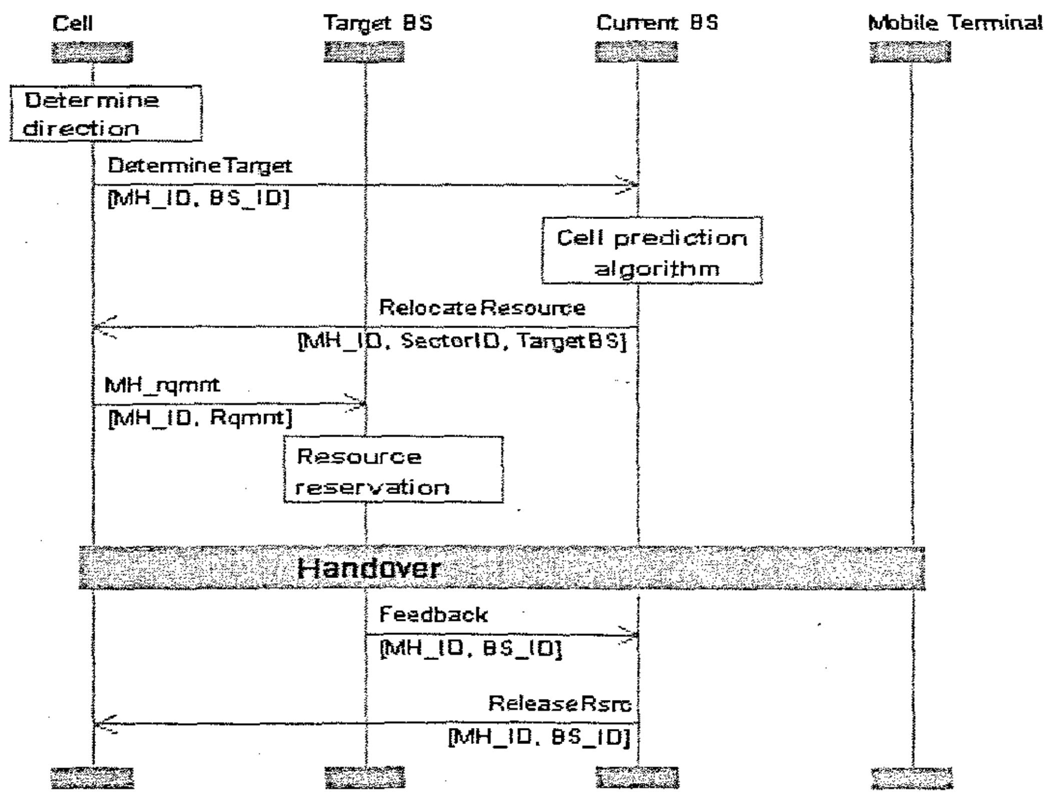

Figure 4.6: NCPS Message Sequence Diagram

At some time after the reservation of resources, the MH executed a handoff operation. The BS to which the MH has actually been handed over notified the previous BS through the **Feedback [MH_ID, BS_ID].** This feedback is essential for the operation of the algorithm. The previous BS notified the previous cell to release previously held resources through **ReleaseResources [[MH_ID, BS_ID]]** signal.

4.9.2 User Interface

The simulation is fully user operated. The user commences the simulation by starting the simulation thread, with the **RunSim** button. Once the simulation is running, the NCPS is illustrated depending on where in a particular cell the active MH is. The real process starts when an MH enters the critical region. This is when the prediction method is invoked, which returns the target BS depending on the current sector and the direction of the MH. The naming of the sectors is done so it could be easy for the prediction method to know the corresponding adjacent cell of each and every sector. The direction of the MH is simulated to be either TOWARDS or AWAY. This is done so as to establish if the MH inside the critical region is arriving or departing. There is a **checkLocation()** method, which continuously check the location of the MH. This method returns ***false*** if the MH is not in the critical region otherwise it returns ***true.***

We also simulates the sending of packets from a corresponding node (CN) so as to trace the number of packets lost due to false prediction. An interesting point worth mentioning is that packets could also get lost due to insufficient available resources in a predicted cell. This aspect is simulated by assigning a random buffer space of a maximum of 100 BU to each BS. If an incoming MH's requirement is greater than the available BUs at each BS, packets directed to that MH are lost. The user interface of the simulation is as shown in figure 4.7

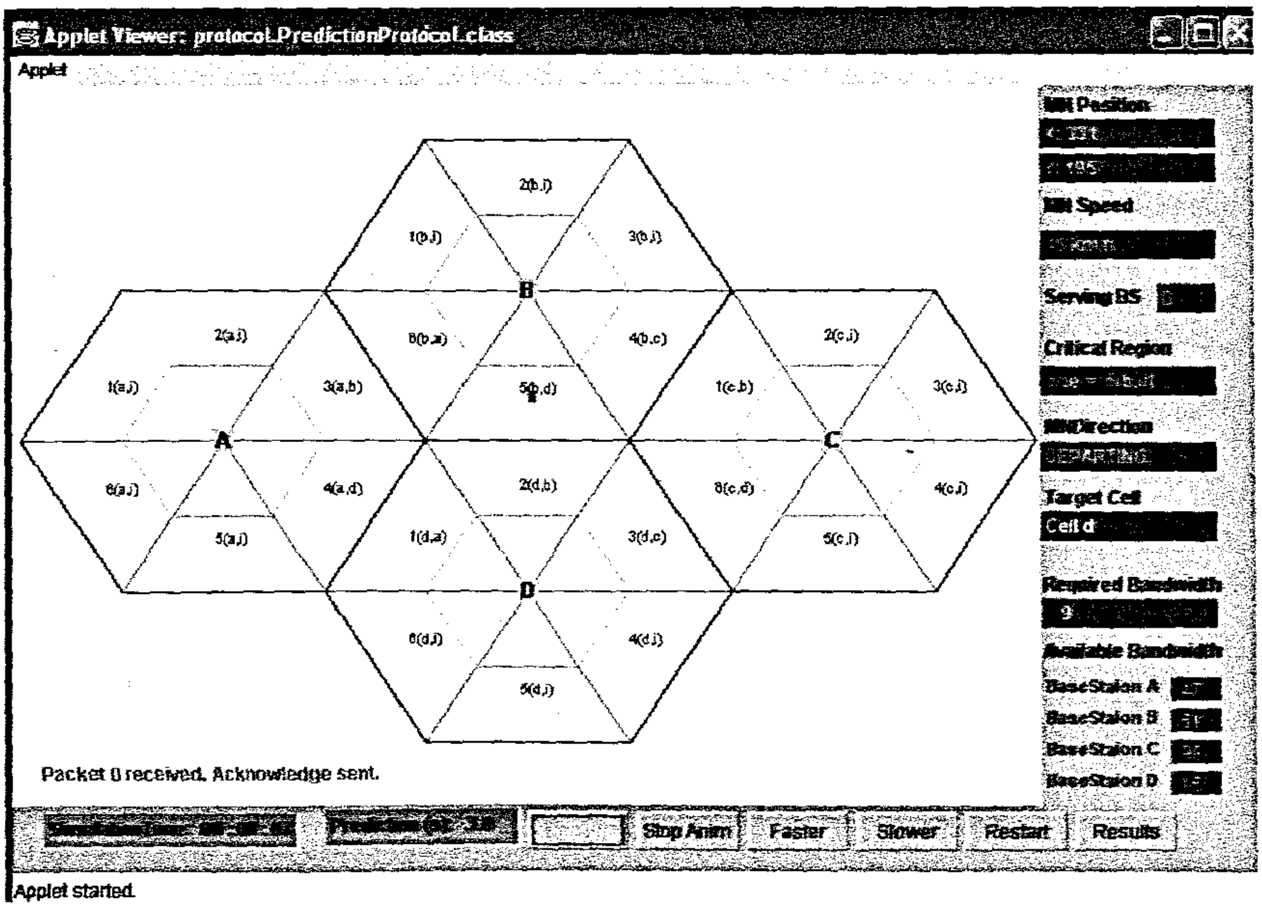

Figure 4.7: NCPS User Interface prototype

The user interactivity of the simulation can be summarised as follows:

1. **RunSim button** for starting the simulation;
2. **Stop/ Resume button** for stopping and resuming the simulation;
3. Simulation speed can be controlled using **Faster/Slower** buttons;
4. The simulation can be restarted anytime using the **Restart button** and
5. Simulation results can be shown by clicking the **Results** button.

4.10 Limitations of the Simulation

The limitations of the simulation are as follows as far as the next-cell prediction scheme is concerned:

1. In reality, the boundary between the critical region and the non-critical region can be dynamically adjusted using handoff requests. In this simulation, we used an appropriate component of the user interface to adjust the critical boundary manually;
2. Only the sender and the receiver were shown. The packets were not shown when they were transmitted and
3. in this simulation only finite number of packets was used. But this is not the case in reality.

CHAPTER FIVE

SIMULATION RESULTS AND ANALYSIS

5.1 Introduction

In this chapter, results of five simulation tests are presented. Tests 1 through test 4 represent the variations of the proposed scheme. While test 5 gives the results of a history-based scheme reported in the literature. A detailed analysis of the results is also presented. It is to be noted that each test is depicted as NCPSx, where x stands for the test number.

5.2 Results and Plots

By incorporating bandwidth reservation into the simulation, two variant schemes were created (this was only done for the purpose of comparing) – one that utilises reserved bandwidth information (NCPS1) and one that does not (NCPS2). We varied the level of user mobility (that is speed). We logged the behaviour of NCPS1 and NCPS2 throughout the maximum of 12 hours (for each simulation run) for a total of 600 predictions and handovers. The scale (m/s for km/h) that we used allowed the model to experience +/-600 total predictions, which we regard as good for the simulation. Figure 5.1 shows how different speeds of the MH affected the accuracy of the NCPS1 (without consideration of resource availability in the predicted cells). The corresponding data for figure 5.1 is presented in Table 5.1.

Table 5.1: Set of data for Accuracy vs. Speed

User's Speed	Time (in hours)											
	1	2	3	4	5	6	7	8	9	10	11	12
Slow Users	0.27	0.58	0.925	0.927	0.930	0.936	0.89	0.91	0.92	0.90	0.92	0.95
Medium Speed	0.25	0.54	0.856	0.87	0.885	0.891	0.86	0.862	0.874	0.87	0.86	0.868
Highly Mobile	0.22	0.512	0.852	0.881	0.866	0.845	0.84	0.835	0.824	0.832	0.838	0.84

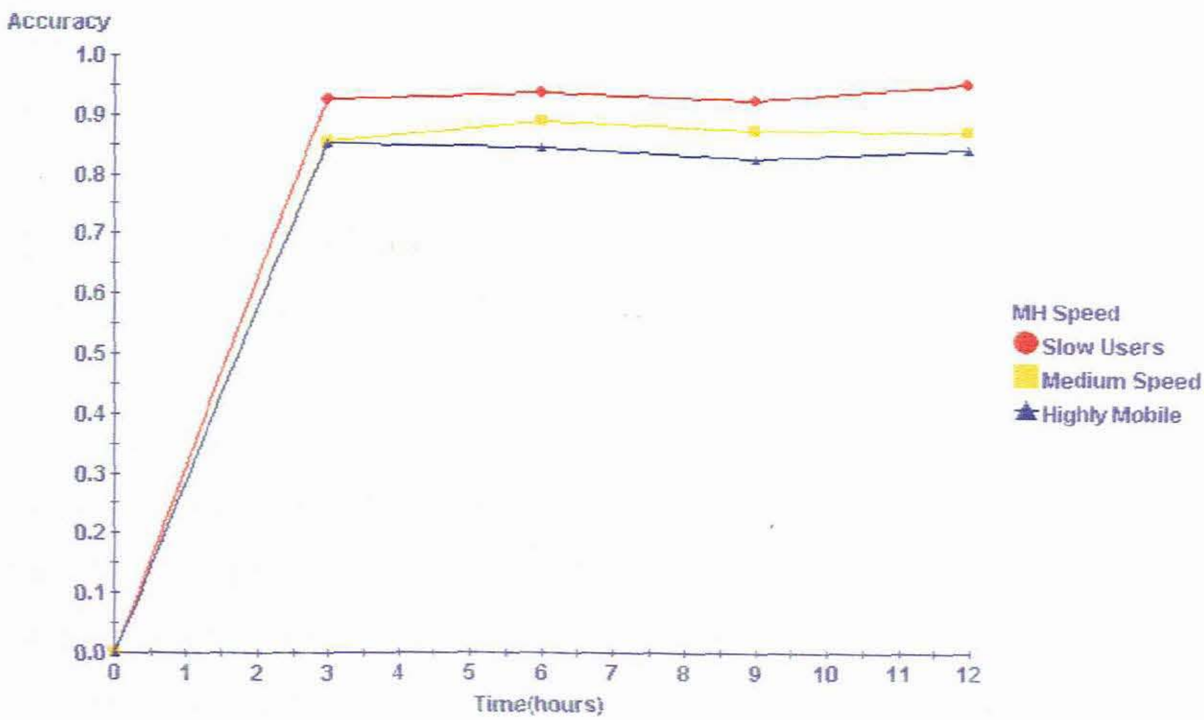

Figure 5.1: Accuracy vs. Speed (resources readily available), NCPS1

The performance of NCPS1 and NCPS2 were evaluated under different speed limits while keeping the CR size constant. Figure 5.1 illustrates the behaviour of the NCPS1 for slow, medium and highly mobile users. The simulation environment allowed users to move in any arbitrary direction while varying their speed in random intervals. We first assumed that the predicted cell would always have enough bandwidth to cater for the incoming MH. Default parameters were set such that when the simulation starts running, the accuracy is 0. It can be seen that as the simulation runs, the graph increases steadily (from time 0 to 2 hours) for all type of users. This is due to the unavailability of CRPs for new systems because no handovers were experienced. A rather stable trend is observed for slow users immediately after 2 and half hours, while there was a small decrease in highly mobile users' trend.

It can be seen that the mean accuracy is 92% for slow users, 87% for medium speed users, and 85% for high-speed users. It needs to be noted that these values were obtained with the assumption that resources are readily available. The scheme performs well when compared to traditional schemes, which do not consider mobile host parameters, such as speed, for their predictions. The advantage here is that predictions are made within parameters of high degree of accuracy for all type of users. We measured accuracy as the ratio of the total number of successful predictions to the total number of executed predictions.

$$Accuracy = \frac{Successfull\ Predictions}{Total\ Predictions} \qquad (5.1)$$

Figure 5.2 shows how different speeds of the MH affected the accuracy of the NCPS2 (with consideration of resource availability). The corresponding data for figure 5.2 is presented in Table 5.2.

Table 5.2: Set of data for Accuracy vs. Speed (limited resources)

User's Speed	Time (in hours)											
	1	2	3	4	5	6	7	8	9	10	11	12
Slow Users	0.26	0.541	0.825	0.86	0.88	0.866	0.85	0.87	0.90	0.882	0.861	0.846
Medium Speed	0.25	0.55	0.856	0.84	0.852	0.851	0.85	0.83	0.844	0.86	0.84	0.854
Highly Mobile	0.267	0.38	0.833	0.85	0.81	0.828	0.80	0.84	0.83	0.85	0.84	0.845

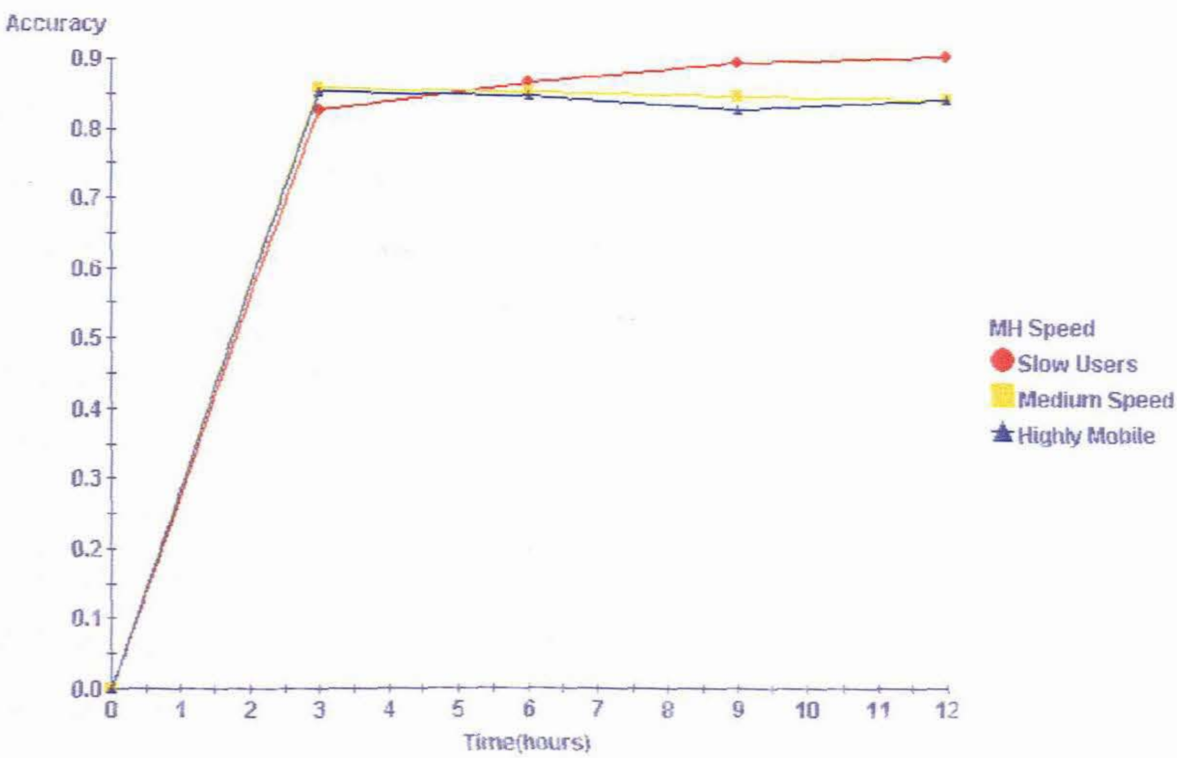

Figure 5.2: Accuracy vs. Speed (limited resources), NCPS2

Figure 5.2 shows the behaviour of the NCPS2 with an assumption that bandwidth is limited during handoff. Again, default values for the accuracies for all users were 0. It can also be seen that as the simulation starts running, the graphs increases steadily (from time 0 to 2 hours) for all type of users This is due to the unavailability of CRPs for new systems because no handovers were experienced. A rather stable trend is observed for slow users immediately after 2 and half hours, while there was a small decrease in highly mobile users' trend. It can be seen that the mean accuracy is 88% for slow users, 84% for medium speed users, and 82% for high-speed users. If we now compare these results with those of NCPS1, we can see that there is not much difference on the behaviour of the schemes. The reason could be that in the sectored-cell approach, unlike in neighbourhood-based schemes, bandwidth is not divided into portions and then reserved in a number of neighbouring cells so as to cater for any possibility of the mobile user heading towards one of those cells. This exercise will always decrease the bandwidth by a very large percentage, which is 5/6 neighbouring cells where resources are stored unnecessarily.

We have also simulated our scheme (NCPS3) under various sizes of CR so as to learn the impact the CR size has on the accuracy of the scheme. Also, the simulation was allowed to run for 12 hours.

Table 5.3: Set of data for Accuracy vs. Critical Region Size

Mobile Speed	Size of the Critical region									
	10	20	30	40	50	60	70	80	90	100
Slow Users	0.0	20.5	48.9	100	100	100	89.0	80.0	80.0	60.0
Highlly Mobile	0.0	0.0	20.5	55.0	100	100	95.2	90.0	78.0	55

Figure 5.3 shows how the size of the CR affects the accuracy of the NCPS. The corresponding data for figure 5.3 is presented in Table 5.3.

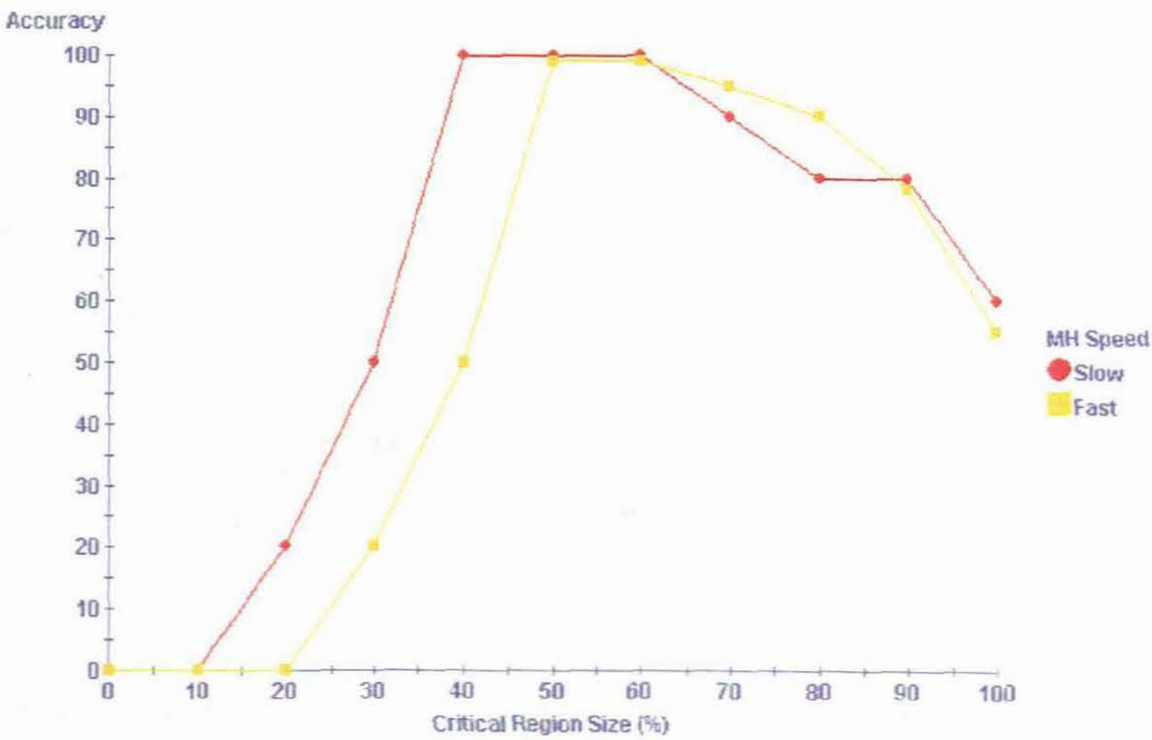

Figure 5.3: Accuracy vs. Critical Region Size, NCPS3

The performance of NCPS3 was evaluated under different sizes of CR for different speed of mobile users. The size of the CR was computed to be inversely proportional to the cell radius. Default accuracy for all users was set at 0%. It can be observed from Figure 5.3 that small values of CR result in more accurate predictions. For instance, 100% accuracy is achievable where the CR is between 35% and 60% of the sector area for slow speed users and between 40% and 60% for high-speed users. For example, the accuracy decreases from 100 to 60 in areas where the CR is very small (i.e., above 60% for both type of users). The explanation for that is that small values of CR result in predictions being made on areas very close to the boundary of the cell. Another reason is that when the mobile user is close to the boundary of the cell and is moving away from the current cell, the most probable outcome is that the user will enter the neighbouring cell it is moving towards. This leads to accurate predictions. However, the graph for highly mobile users suggests that predictions for high-speed users should not be made very close to the cell boundary because there is not enough time left to find enough resources to accommodate this user before the user enters the coverage area of the predicted cell.

Therefore when deciding upon CR values for mobile users with varying speeds, the above discussion should be taken into consideration. CR values that give reasonable prediction accuracy while allowing enough time to find enough resources to handover users in the target cells should be chosen. Table 5.6 reflects the values selected to satisfy both criteria.

Table 5.6: Recommended CR Values

Users	CR Values (%)
Slow	Between 35 & 60
High-speed	Between 40 & 60

Another interesting point worth mentioning is that when CR values are above 60%, it means that the mobile user is near the centre of the cell; hence there is more freedom to change path before entering another cell. This tends to make it difficult to determine to which cell the user is heading. The graphs therefore show that the predictions of the next-cell an MH will move to, is less accurate near the centre of the cell. Users that are near the centre of the cell, will normally take a longer period to reach any of the neighbouring cells.

Table 5.4: Set of data for Next-Cell Prediction Efficiency

Predictions	Time (in hours)							
	0	1	2	3	4	5	6	7
Accurate Predictions	80.0	85.0	90.2	88.0	85.0	78.8	85.6	88.0
Missed Predictions	0.0	5.0	12.0	9.0	14.0	12.2	9.6	8.4

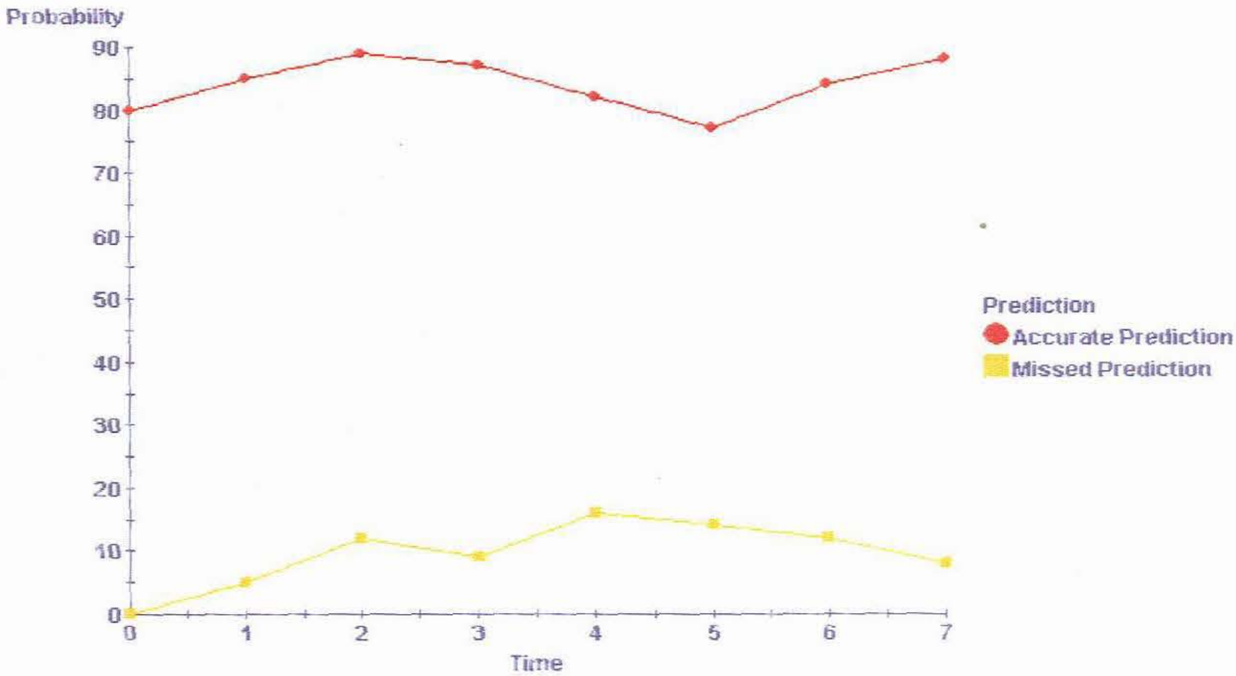

Figure 5.4: Next-Cell Prediction Efficiency, NCPS4

Finally, we simulate probability that a predicted cell is indeed the cell an MH has been handed over (that is NCPS4).Figure 5.4 shows the probability that the NCPS output (which is the target cell) is the one to which the terminal has really been handed over to, i.e., accurate prediction. The corresponding data for figure 5.4 is presented in Table 5.4.

We also observed the performance of NCPS4, which is the proposed scheme. The above figure shows that the accuracy of the scheme remains high, despite the simulation time. Accurate predictions are averaging at 80%, while missed prediction remains below 20%. The explanation for this is that the proposed scheme is not dependent on mobile user history for making predictions. Thus, it does not have to wait until some data is available so it could start making accurate predictions. Another explanation is that, the proposed scheme does not predict all neighbouring cells for resource reservation, it predicts only 1/6 of these, which is an efficiency of 80%.

The results of the above plot are compared to the results of the Proxies + Path Prediction plots (we shall call it NCPS5) given in the literature [25]. Figure 5.5 presents the results of the Proxies + Path Prediction scheme. Table 5.5 presents the corresponding set of data.

Table 5.5: Set of data for Path Prediction Efficiency

Predictions	Time (in days)						
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Perfect Hit	30.0	25.0	29.0	47.0	52.0	47.0	49.0
Miss	70.0	45.0	32.0	29.0	26.0	34.0	32.0
Perfect Hit + Hit to Second best	55.0	55.0	68.0	70.0	75.0	67.0	69.0
Hit to Second Best	35.0	35.0	40.0	25.0	26.0	22.0	20.0

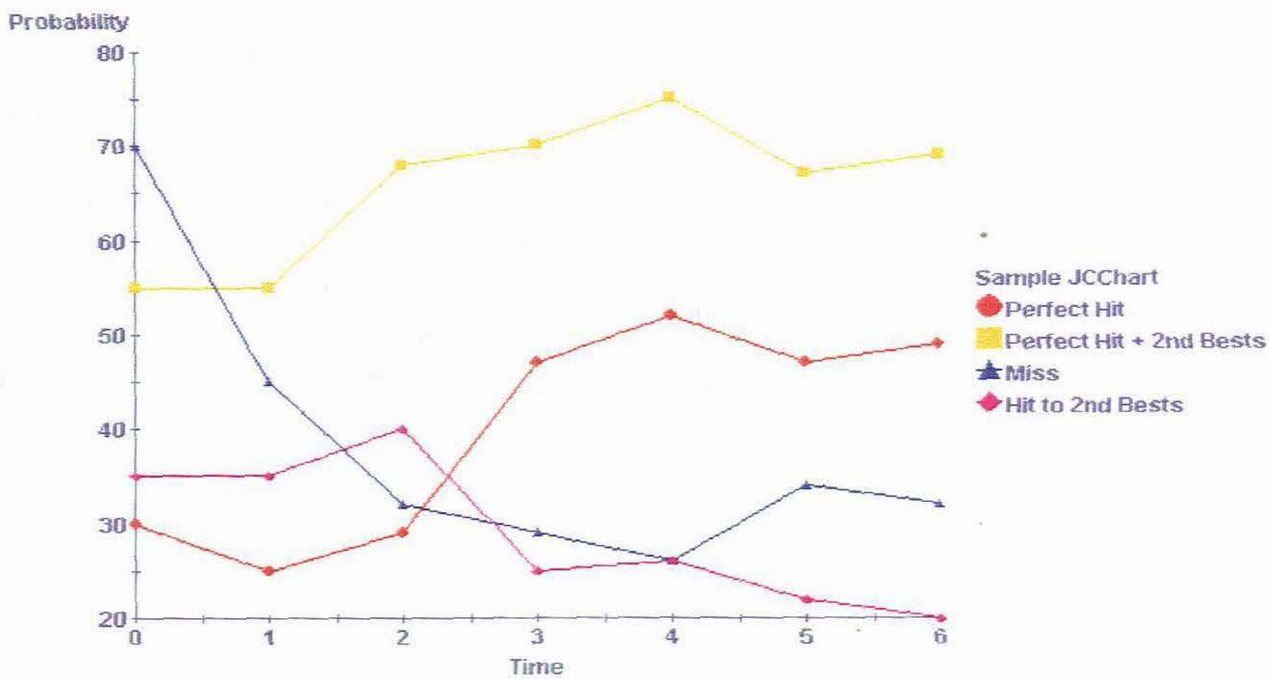

Figure 5.5: Path Prediction Efficiency, NCPS5

Finally, we considered the performance of NCPS5, which is claimed to be the best history-based scheme in the literature. In this scheme, the series for missed prediction shows that the prediction misses for their scheme is unreasonably high (70%) for the first days of the simulation. It is only after the third day of the simulation that the average number of inaccurate predictions becomes reasonable around 32%. The graph for their Perfect Hit remains below 30% for the first three days, but as soon as user mobility patterns become known to the system, then their algorithm starts behaving within parameters of acceptable behaviour, between 50 and 60 percent. This would be due to the fact that all history-based schemes are likely to be able to make accurate predictions for first time users. This is because they rely heavily on user data for their prediction. Results presented in Figure 5.4 prove that to achieve a high degree of accuracy (90%), one does do not really need to keep user mobility history. The series for missed call is stable around 7%, which we think is acceptable for any prediction scheme. The series for accurate prediction is also averaging around 90%, which is good as well.

The NCPS4 has better accuracy than NCPS5. These demonstrate that mobility prediction schemes based on mobile parameters are more accurate, thus leading to more efficient reservations. In terms of comparison, the simulation results of the Proxies + Path Prediction algorithm, shows a performance quite close to 68% (for their Perfect Hit + Hit to second best) for a simulation period of 1 week. This is the performance that our scheme achieved immediately the algorithm started executing. Secondly, we did not have to keep user profiles for our scheme to give better predictions. Keeping mobility profiles in the system leads to low utilisation of the system resources since many overheads are introduced. Therefore, our results strongly suggest that significant improvement in reservation efficiency may result from the use of mobility predictions based on mobile parameters.

Another important point worth mentioning is that neighbourhood-based schemes were not simulated here since there is no effort made in predicting the next cell in these schemes. However, resources are reserved in all cells surrounding the current cell irrespective of the mobile users' whereabouts.

CHAPTER SIX

SUMMARY AND CONCLUSION

6.1 Summary

Provisioning of QoS and efficient reservation of resources is fundamental in mobile wireless networks. With the advent of real-time data networks and the increasing need for seamless mobility, efficient resource reservation techniques and fast handover schemes are essential.

In this research work, a next-cell prediction scheme was presented. This scheme partitioned a radio cell into six sectors to accurately predict the next-cell a mobile node would move to. Partitioning each radio cell into small sectors reduced the entire tracking area, which increased the positioning accuracy.

It has been shown to this effect with the help of simulation results, that a high level of accuracy in prediction can be achieved with considerable decrease in the total tracking area. The behaviour of the NCPS was tested under different conditions. Here, two parameters that are highly capable of affecting the accuracy of the NCPS were identified: the speed of the MH and the size of the critical region. The simulation ran under four tests. Test 1 demonstrates the behaviour of the NCPS1 with an assumption that resources are readily available. The speed of the MH was varied while the size of the critical region remained constant. The results showed that the speed of the mobile host affects the accuracy of the prediction. It was also observed that when the scheme was simulated without consideration of resource reservation, a high level of accuracy is achieved. Test 2 demonstrates the behaviour of the NCPS2 with resource reservations. The speed of the MH was varied while the size of the critical region remained constant. Observations from this test were similar as in test 1 with respect to mobility speeds except that the considerations of resources affected the performance of the scheme a great deal. This means that for any prediction scheme to be complete, it should cater for the possibility of limited resources in the predicted cell.

Test 3 demonstrates the behaviour of the NCPS3 under various sizes of critical regions for different speed of mobile users. Test 4 demonstrates the probability that the NCPS4 output, which is the target cell, is the one to which the MH has really been handed over to. The results of this test were compared with the results of [25].

With a consideration of MH's real-time parameters, such as speed, location and direction, we ascertain that accurate predictions can be achieved. While an MH may change its direction and speed as it approaches the cell boundary, the target cell prediction was erroneous only if the trajectory of the MH changed so much that it entered a different cell from the one previously predicted. A precise knowledge of the exact time at which a handoff-request was to be made was not required, which is an intractable task since a handoff occurred anywhere within the CR, and a MH's speed and direction also changed randomly. Predictions were made periodically and alternative reservations were attempted if the previous decision has become invalid. By regularly updating the CRPs, the scheme is robust against changes in handoff policies, as well as changes in terrain and manmade features that could affect radio propagation.

6.2 Conclusion

The NCPS4 has better accuracy than NCPS5. These demonstrate that mobility prediction schemes based on mobile parameters are more accurate, thus leading to more efficient reservations. The simulation results of the Proxies + Path Prediction algorithm, show a performance quite close to 68% (for their Perfect Hit + Hit to second best) for a simulation period of 1 week. This is the performance that our scheme achieved immediately the algorithm started executing. We did not have to keep user profiles for our scheme to give better predictions. Keeping mobility profiles in the system leads to low utilisation of the system resources since much overhead is introduced. Therefore, the results of the NCPS strongly suggest that there is significant improvement in reservation efficiency that may result from the use of mobility predictions based on mobile parameters.

6.3 Future Work

The mobility prediction scheme presented here focuses only on next-cell prediction. The scheme can also be extended to include features that could simplify the way handover needs to be done in a sectored-cell structure. The simulation for the scheme was run without any consideration of objects, like mountains and tall buildings, which could obstruct the view in the network. For a strong realistic feel, we recommend extension of the scheme to accommodate that. The work presented in this dissertation is focused on mobility prediction in wireless mobile networks, and this is probably the most fundamental challenge in mobile ad hoc networks. However, we are well aware that mobile ad hoc networks do not have any fixed communication infrastructure. Exploring a way to extend the proposed scheme to also accommodate mobile ad hoc networks is recommended.

REFERENCES

[1] Acampora, A. and Naghshineh, M., "An architecture and Methodology for Mobile-Executed Handoff in Cellular ATM Networks," *IEEE JSAC,* 12(8), October 1994, pp. 1365-1374.

[2] Soh, W-S. and Kim, H. S., "QoS Provisioning in Cellular Networks Based on Mobility Prediction Techniques," *IEEE Commun. Mag., pp. 86 –92, Jan. 2003.*

[3] Levine, D.A., Akyildiz, I.F., and Naghshineh, M., "A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks using the Shadow Cluster Concept", *IEEE/ACM Trans. Networking,* vol. 5(1), Feb. 1997, pp. 112.

[4] Sunghyun, C., "Adaptive Bandwidth Reservation and Admission Control in QoS-Sensetive Cellular Networks," *IEEE Trans. Distr. Syst*, Vol. 13(9), Sept. 2002.

[5] Riera, M. and Aspas, J., "Variable Channel Reservation Mechanism for Wireless Networks with Mixed types of Mobility Platforms," *in Proc. IEEE VTC'98*, May 1998.

[6] Oliveira, C., Kim, J.B., and Suda, T., "Adaptive Bandwidth Reservation Scheme for High-speed Multimedia Wireless Networks," *IEEE J. Select. Areas Commun.*, vol. 16(6), Aug. 1998, pp. 858874.

[7] Chua, K. C. and Choo, S. Y., "Probabilistic Channel Reservation Scheme for Mobile Pico/ Microcellular Networks," *IEEE Communication Letters, 2(7),* July 1998, pp. 195-196.

[8] Liu, T., Bahl, P. and Chlamtac, I., "Mobility Modeling, Location Tracking, and Trajectory Prediction in Wireless ATM Networks," *IEEE JAC, 16(6)*, August 1998, pp. 922-936.

[9] Chan, J., De Silva, R., Zhou, S. and Seneviratne, A., "A Framework for Mobile Wireless Networks with an Adaptive QoS Capability," in *Proc. MoMuC'98*, Oct 1998.

[10] Luo, X., Thng, I. and Zhuang, W., "A Dynamic Pre-Reservation Scheme for Handoffs with QoS Guarantee in Mobile Networks," *Proc. IEEE Int'l Symp. Computers and Comm.*, July 1999.

[11] Jannick, J. *et al*, "Efficient and Flexible Location Management Techniques for Wireless Communication System," *ACM/Balzer Wireless Networks*, 3(5), 1997, pp.361-374.

[12] Krishna, P., Vaidya, N. and Pradhan, D., "Static and Adaptive Location Management in Mobile Wireless Networks," *Computer Communications*, 19(4), Apr. 1996, pp.321-334.

[13] Bharghavan, V. *et al*, "The TIMELY Adaptive Resource Management Architecture," *IEEE Personal Communication Magazine*, August 1998, pp. 20-31.

[14] Bless, R. et al, "Quality of Service Signalling in Wireless Ip-Based Mobile Networks," in *Proc. Of VCT2003-fall, Orlando*, FL, USA, October 2003.

[15] Chin, M. H. and Bassioni, M.A., "Predictive Schemes for Handoff Prioritisation in Cellular Networks based on Mobile Positioning," *IEEE J. Select. Areas Commun., Vol. 18(3), March 2000, pp. 510 – 522.*

[16] Talukdar, A., et al, "MRSVP: A Resource Reservation Protocol for an Integrated Services Network with Mobile Hosts," *ACM/WINET*, May 1999.

[17] Pajares, A., et al, "An Approach to Support Mobile QoS in an Integrated Services Packet Network, " *in Proc. IQWiM Workshop*, Apr. 1999.

[18] Awduche, D. and Agu, E., "Mobile Extentions to RSVP," *in Proc. IEEE ICCN'97*, 1997.

[19] Choi, S. and Shin, K. G., "Predictive and Adaptive Bandwidth Reservation for Handoffs in QoS-Sensitive Cellular Networks", in *Proc. ACM SIGCOMM98*, Vancouver, British Columbia, Sep. 1998, pp. 155-166.

[20] Bharghavan, V., and Jayanth, M., "Profile-Based Next-Cell Prediction in Indoor Wireless LAN," in Proc. *IEEE SICON'97*, Apr. 1997, available at http://shiva.crhc.uiuc.edu/publications.html

[21] Erbas, F. *et al*, "A Regular Path Recognition Method and Prediction of User Movements in Wireless Networks," *Vehicular Technology Conference VTCF01*, Atlantic City, 2001.

[22] Lui, G and Maguire, G. Jr., "A Class of Mobile Motion Prediction Algorithms for Wireless Mobile Computing and Communications," *ACM/Balzer MONET*, 1(2), 1996, pp. 113-121

[23] Chan, J. and Senevirathe, A., "A Practical User Mobility Prediction Algorithm for Supporting Adaptive QoS in Wireless Networks" *IEEE International Conference In Networks*, Sept. 1999.

[24] Jayasuriya, A. and Asenstorter, J., "Mobility Prediction Model for Cellular Networks based on Observed Traffic Patterns", Available at

http://www.itr.unisa.edu.au/~aruna/papers/paper356-235.pdf

[25] Hadjiefthymaides, S. and Merakos, L., "Proxies + Path Prediction: Improving Web Services Provision in Wireless Mobile Communication". 2000, Available at http://citeseer.ist.psu.ude/550402.html

[26] Majoor, R., "Quality of Service in the Internet Age," in *SATNAC2003* proceedings, pp. 1-4, *Sept, 2003.*

[27] Bochmann, G.V., Hafid, A., "Some Principles for QoS Management", *Distributed, Systems Engineering, vol. 4, pp. 16-27,* IOP Publishing, 1997.

[28] Perkins, C., "Mobile IP", *IEEE Communication Magazine*, May 1997.

[29] Srivastava, M., Mishra, P.P., "On QoS in Mobile Wireless Networks", *Proc. IEEE 7th Int' Wksp on Network and Operating Systems Support for Digital Audio and Vedio*, pp. 147 – 158, 1997.

[30] Perkins, C.E., "IP Encapsulation within IP," *Network Working Group,* Request for Comments, 2003, October 1996.

[31] Perkins, C.E., "IP Mobility Support," *Network Working Group,* Request for Comments, 2002, October 1996

[32] Perkins, C.E. and Johnson, D.B., "Route Optimization in Mobile IP," *Mobile IP Working Group*, Internet Draft – work in progress, November 1997.

[33] Gurumoorty, K.K., "Capacity Analysis of CDMA Systems", Available at http://www.itc.ukans.edu/~krishk/Documents/CDMA.pdf.

[34] Nguyen, A.V., Wan, P. and Frieder, O., "Channel Alternation And Rotation For Tri-sectored Cellular Systems" available at ttp://journals.kluweronline.com/article.asp?PIPS=5119365.

[35] Freedman, A. and Hadad, Z., "Handoff Schemes Overview and Guidelines for Handoff Procedures in 802.16," *IEEE 802.16 Presentation sub. Template (Rev. 8.3)*, Sept. 2002.

[36] Hansen, H. et al, "Description of the Handover Algorithm for the Wireless ATM Network Demonstrator (WAND)", *in Proc. ACS Mobile Communications Summit*, Granada, Spain, November 27-29, 1996.

[37] Zeng, Q-A. and Agrawal, D. P., "Handoff in Wireless Mobile Networks," 2001, available at *http://media.wiley.com/product-data/*excerpt/28/04714190/04714190280.pdf.

[38] Grillo, D., Skoog, R. A., Chia, S. and Leung, K. K.," Teletrac Engineering for Mobile Personal Communications" *in ITU-T work: the need to match practice and theory*, IEEE Pres. Commun., pp. 38-58, Dec. 1998.

[39] Dru, M-A. and Saada, S., "Location–based Mobile Services: the Essentials", 2001, available at: *www.alcatel.com/doctyps/*articlepaperlibrary/pdf/ART2001Q1/gb/14drugb.pdf.

[40] Yi-Bing, L. and Victor, W.M., "Eliminating Boundary Effects of a large-scale Personal Communication Service Network Simulation", *in: ACM TOMACS*, vol 4(2), Apr. 94.

[41] Omar, S. and Rizos, C., "Incorporating GPS into Wireless Networks: Issues & Challenges", in SatNav, pp. 1-7, July 2003.